University of Alberta

Systematics of Western North American *Dioryctria* (Lepidoptera: Pyralidae)

by

Amanda Diane Roe

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

In

Systematics and Evolution

Department of Biological Sciences

Edmonton, Alberta, Canada

Fall 2006

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-23102-9
Our file Notre référence
ISBN: 978-0-494-23102-9

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Species of *Dioryctria* Zeller (Lepidoptera: Pyralidae) are notoriously difficult to identify due to variable forewing morphology, few species specific genitalic features, and broad sympatric distributions. Accurate species delimitation and identification are essential for effective monitoring and control of this conifer pest. I explored several aspects of species delimitation in *Dioryctria* using a combination of molecular, morphological, ecological, and behavioural characters. I compared mtDNA variation, adult morphology, larval host association and pheromone attraction in a complex of *Dioryctria* species and identified eight distinct mitochondrial DNA (mtDNA) lineages that corresponded to eight *Dioryctria* species. Morphology, pheromone attraction, and larval host association also delimited these species and provided useful diagnostic characters. To evaluate how patterns of molecular evolution affect species delimitations I examined mitochondrial cytochrome c oxidase I and II (COI-COII) nucleotide substitution patterns within and between sister species of Lepidoptera and Diptera. I found heterogenous accumulation of maximum divergence and phylogenetic signal, overlap between intra- and interspecific divergence, and no optimally informative 600 bp location (length chosen to assess the information content of the DNA barcoding). I used a combination of independent molecular loci, forewing morphology, geography, and larval host association to test the traditional delimitation of two sympatric *Dioryctria* species, *D. pseudotsugella* and *D. reniculelloides*. I found restricted gene flow between an eastern *D. reniculelloides* clade and a western *D. pseudotsugella* clade. Two independent molecular loci, morphological variation, and larval host association supported these two clades. However, diagnosis of these two species remains difficult

variation and future examinations are needed to elucidate which barriers to gene flow are maintaining this sympatric distribution. To improve species delimitations I recommend that researchers 1) combine multiple lines of evidence to validate, and cross validate species boundaries, 2) broadly sample biological and geographic variation within and between closely related species, and 3) maximize mtDNA sequence length to increase the probability of sampling regions of high divergence, minimize stochastic variance in estimating total COI-COII divergence, and incorporate regions of informative phylogenetic signal.

Acknowledgments

It has taken nearly five years to reach this point, and there have been many people who have helped me along the way. First of all, I wish to thank Felix, my supervisor, for enthusiastically introducing me to systematics and molecular evolution, and for allowing me to pursue my entomological epiphany. Without his guidance and encouragement, I would have never known the fascinating world of *Dioryctria* or DNA.

I would also like to thank my committee members, Heather Proctor and John Spence, for all their suggestions and insights. Heather was an invaluable resource for statistics and ordinations and she encouraged me to explore an avenue of research that would have otherwise fallen by the wayside. John was invaluable for asking the difficult questions and always reminding me about big picture.

My project was funded from both a USDA contract and NSERC Discovery Grant awarded to F. A. H. Sperling. I was also supported by two NSERC postgraduate fellowships I received over the course of this degree.

I must express my gratitude to all the collectors, forest managers, and museum curators who have collected specimens or provided museum material for this thesis, for this research would have been impossible without their contributions. I also greatly appreciate the efforts of Amber and Lisa for laboratory work, larval rearing and wonderful company while gallivanting around the countryside. Of course, I cannot forget to thank Gary Grant, who not only provided access to pheromone trapped *Dioryctria* specimens, but taught me about the kinds of intricacies I may face in academia.

I would also like to thank the members of the Sperling lab and the other members of the entomology grad lounge for animated discussions about life, the universe and everything. Heather, Keith, Jeff, Tara and Claire have been such wonderful friends and fantastic researchers. I feel honored to have spent so much time with them throughout the years.

I must also thank my family for being with me through it all. They have put up with all my eccentricities, and even the occasional escaped snake (sorry Mom!). Without their love and support, I know I would not have made it to where I am now, or fully appreciate the world around me.

Finally, I must thank Chris for a great many things. Thank-you for all of our discussions; statistical, biological, philosophical and everything in between. Thank-you for all your support, understanding and encouragement. You are always there when I need you the most and coached me through the difficult times. I only hope that I can do as much for you as you have done for me. Thank-you.

Table of Contents

Chapter 1: General Introduction 1

The Genus *Dioryctria* 1

Diagnostic Characters 2

Molecular Markers 3

Morphological and Ecological Characters 4

Scope of the Thesis 4

Literature Cited 7

Chapter 2: Identification of *Dioryctria* (Lepidoptera: Pyralidae) in a seed orchard at Chico, California 11

Introduction 11

Materials and Methods 13

Collection Sites and Species 13

Collection Methods 14

Molecular Techniques 15

Phylogenetic Analyses of DNA Sequence 16

Morphological Techniques 16

Results 18

Analysis of 475 bp COI Fragment 18

Morphological Identifications. 19

Larval Host Plant Associations 21

Pheromone Analysis 22

Phylogenetic Analysis of 2.3 kb fragment 22

Information Content of Sequence Fragments 23

Discussion 24

MtDNA Lineages 24

Morphological Identification 26

Reared Material and Pheromone Attraction 27

Phylogenetic Relationships 29

Literature Cited 44

Chapter 3: Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for species recognition.....53

Introduction.....53

Methods.....56

Overview.....56

Available Material.....57

Sliding Window Analysis.....58

Randomizations.....59

Localized Nucleotide Divergence as an Indicator of Total COI-COII Divergence.....60

Fragment Length.....60

Phylogenetic Signal.....61

Results.....62

Intraspecific Diversity.....62

Interspecific Divergence.....63

Overlap of Intra- and Interspecific Divergence.....64

Randomizations.....65

Chimeric Sequences.....65

Localized Divergence Relative to Total COI-COII Divergence.....66

Fragment Length.....66

Phylogenetic Signal.....67

Discussion.....68

Mutation Hot Spots.....68

Intra- verses Interspecific Divergence.....69

Fragment Size and Location.....70

Phylogenetic Signal.....72

Conclusion.....73

Literature Cited.....85

Chapter 4: Delimitation of two sympatric *Dioryctria* species using an integrative taxonomic approach.....94

Introduction.....94

Materials and Methods.....96

Species and Study Area.....96

Laboratory Methods.....97

Genetic Analyses *98*
Morphometric Measurements *100*
Morphometric Analyses *100*
Hybrid Index *101*

Results **102**
Sequence Summary Statistics *102*
mtDNA 102
Nuclear 102
Phylogenetic and Population Genetic Analyses *103*
Parsimony 103
Nested Clade Analysis 103
Genetic Differentiation 103
Morphometric Analysis *104*
Hybrid Index *105*

Discussion **106**

Literature Cited **123**

Chapter 5: General Conclusions **128**

Sampling Bias **129**

Incongruence **129**
Incomplete Lineage Sorting *129*
Introgression and Interspecific Hybridization *130*
Nuclear Pseudogenes *131*

Recommendations **131**

Future Directions **132**

Literature Cited **134**

Appendix 1 ..... **140**

Appendix 2 ..... **144**

List of Tables

Table 2- 1: Locality and collection information for *Dioryctria* specimens used in mtDNA surveys in northern California, Oregon and British Columbia. 32

Table 2- 2: Mitochondrial DNA primers used in surveying *Dioryctria* spp. over 2.3 kb of COI-COII. 35

Table 2- 3: Wing phenotypes for *Dioryctria* specimens used in mtDNA survey. 36

Table 3- 1: Intraspecific nucleotide diversity (%) for 23 species with multiple specimens (N), and are calculated as the average diversity between all specimens within a species. Full COI-COII diversity (Total), maximum diversity (Max, 600 bp fragment), and minimum diversity (Min, 600 bp fragment) are shown. Midpoints are shown in *D. yakuba* numbers, and regions with multiple equal divergence values were combined when less than 100 bp apart. Diversity was also examined for representative 600 bp fragments for the DNA barcoding region (BC, midpoint *Dy*# 1856), Jerry-Pat (J-P, midpoint *Dy*# 2599), George-Marilyn (G-M, midpoint *Dy*# 3090), and Pierre-Eva (P-E, midpoint *Dy*# 3434). 75

Table 3- 2: Interspecific nucleotide divergence (%) for 23 sister species pairs. Values for full 2.3 kb of COI-COII divergence (Total), maximum divergence (Max), and minimum divergence (Min) are shown. Midpoints are shown in *D. yakuba* numbers, and regions with multiple equal divergence values were combined when less than 100 bp apart. Divergence was also examined for representative 600 bp fragments for the four primer regions described in Table 1. 76

Table 3- 3: Nucleotide divergence and relative percentages of total COI-COII divergence sampled by PCR fragments commonly used in insect systematics. Mean divergence and relative divergence are calculated from 23 sister species pairs. 77

Table 4- 1: Localities for specimens surveyed for molecular and morphometric characters. Latitude (Lat) and longitude (Long) are given in decimal degrees. Number of specimens (N) and rearing information (Host) are also shown. 110

Table 4- 2: Sample locality (for abbreviations, see Table 4-1), sample size (N), haplotype and genotype information for specimens surveyed in this study. Numbers in brackets indicate number of individuals for haplotypes or genotypes, if there are multiple specimens per locality. **111**

Table 4- 3: Nucleotide variation and genotype designation for 29 individuals sequenced for ITS2. For locality (LC) abbreviations, see Table 4-1. Invariant positions are indicated with an apostrophe ('), and gaps in the alignment are indicated with a dash (-). For reference, COI haplotypes (COI Hap) and the location of each specimen in the two main mtDNA clades is indicated (NCA). The specimen in bold highlights the presence of a discordant combination of haplotypes and genotypes. **112**

Table 4- 4: Nucleotide variation and genotype designation for 33 individuals sequenced for EF1α. For locality (LC) abbreviations, see Table 4-1. Nucleotide position is shown relative to *D. melanogaster* (Hovemann et al 1988). Invariant positions are indicated with an apostrophe ('). **113**

Table 4- 5: SAMOVA results for genetic differentiation between three optimal groups roughly corresponding to mtDNA network clades 2-1+2-2, 2-3, and 3-2..... **114**

Table 4- 6: Distribution of four morphometric measurements in the two main mtDNA clades found in the nested clade analysis. All measurements were taken in ImageJ. A comparison of the morphometric character means was conducted using a two-tailed t-test. **115**

Table 4- 7: Morphometric and larval host characters used for scoring a hybrid index for individuals identified as *D. reniculelloides* and *D. pseudotsugella* in mtDNA clades or previous taxonomic identifications. Negative values were assigned to characters that were *D. pseudotsugella*-like, and conversely positive for *D. reniculelloides*-like characters based on the original descriptions and diagnoses of the two species..... **116**

Table 4- 8: Frequency of correct and incorrect identifications using two hybrid indices: 1) morphometric characters only, or 2) morphometric characters and larval host association. Indices were evaluated against specimens classified on the basis of mtDNA clade, and museum specimens identified prior to this investigation. A negative score indicates *D. pseudotsugella*-like characters, while a positive score

indicates *D. reniculelloides*-like characters (for character descriptions see Table 4-7).....117

List of Figures

Figure 1- 1: Common western North American *Dioryctria* species. A) *D. pentictonella*; B) *D. cambiicola*; C) *D. fordi*; D) *D. rossi*; E) *D. pseudotsugella*; F) *D. okanaganella* 6

Figure 2- 1: Maximum parsimony analysis of 475 bp COI. (A) Phylogram of one of 178 most-parsimonious trees (length = 182 steps; CI = 0.692; RI = 0.862) showing phyletic branch lengths. Lineages labeled 1 – 8 are discussed in text. (B) Strict consensus of 178 most-parsimonious trees, with bootstrap values >50% shown above branches and Bremer support values below branches. Haplotype codes are explained in Table 2-1. 38

Figure 2- 2: Flight period for *D. pentictonella* based on all pheromone and light-trapped material included in study, including 1998 pheromone-trapped material. 39

Figure 2- 3: Wing phenotypes collected during three flight periods of *D. pentictonella.* Dark horizontal bars separate number of specimens collected by pheromone traps (below) from other methods (above). Wing phenotype coding is discussed in text. 40

Figure 2- 4: Flight times for specimens collected from three pheromone lures (I, III, V; lure components explained in Table 2-4) from 2000 and 2001 which were also sequenced for mtDNA. Lures with catch totals of fewer than three specimens are not shown. 41

Figure 2- 5: Phylogram of single most parsimonious tree for 2.3 kb COI-COII (length = 840; CI = 0.714; RI = 0.746). Lineages labeled 1 – 8, and Nodes A – G are discussed in text. Known host associations and pheromone attraction are indicated for each species, Bootstrap values >50% are shown above branches and Bremer support values below branches. 42

Figure 2- 6: Comparison of uncorrected sequence divergences within species, within species groups and between species groups for *Dioryctria* specimens sequenced across the full 2.3 kb of COI+COII. Abbreviations of between species groups comparisons are as follows: *zimm* Gr.= *zimmermani* Gr; *pond* Gr.= *ponderosae* Gr. 43

Figure 3- 1: Overview of COI-COII genes, with nucleotide locations as in *D. yakuba* (Clary and Wolstenholme 1985). Structural regions are shown along the upper bar and abbreviated according to Lunt et al. (1996). Commonly used COI-COII regions in insect systematics (Caterino et al. 2000) are identified by their forward and reverse primers (Bogdanowicz et al. 1993, Folmer et al. 1994, Simon et al. 1994). **78**

Figure 3- 2: Sliding window profiles of COI-COII (600 bp window; 5 bp step) of intraspecific nucleotide diversity for 12 of 23 species. Average nucleotide diversity for total COI-COII follows each species name. Black circles indicate DNA barcoding midpoint. Scales on Y-axes differ between upper and lower panels.....**79**

Figure 3- 3: Sliding window profiles of COI-COII (600 bp window; 5 bp step) of nucleotide divergence for 15 of 23 sister species pairs. Total COI-COII nucleotide divergence follows each species name. Black circles indicate DNA barcoding midpoint. Scales on Y-axes are variable.....**80**

Figure 3- 4: Sliding window profiles of COI-COII (600 bp window; 5 bp step) of intraspecific nucleotide diversity (*A. gambiae*) and interspecific divergence (*A. gambiae* x *A. arabiensis*) showing regions of overlap. Black circle indicates DNA barcoding midpoint.....**81**

Figure 3- 5: Sliding window profile of *P. troilus* x *P. palamedes* and profiles of 50 randomly generated sequence pairs. Original profile (thick bold line), mean random divergence (thin bold line) and ± 95% CI (shaded region) of random profiles are shown.....**82**

Figure 3- 6: Sliding window profile of mean relative divergence (bold) and ± 1 standard deviation (shaded) for all 23 sister species pairs. Individual profiles (grey lines) show individual sister species pairs. Midpoint locations of commonly used COI-COII PCR regions and the "optimal" 600 bp region are shown. **83**

Figure 3- 7: Mean relative nucleotide divergence and variance (1 standard deviation) for fragments of increasing length starting at either the 5' end of COI, the midpoint of COI-COII, or the 3' end of COII for 23 sister species. Fragment length started at 200 bp, and increased in 200 bp increments until the entire sequence was included**84**

Figure 3- 8: Phylogenetic time series for *Papilio machaon oregonius*, *Papilio canadensis,* and *Dioryctria zimmermani* with phylogenetic series arranged in

columns and similar total divergences arranged in rows. Each species is compared to increasingly more diverged congeners of roughly equal increments (where available), with total divergence indicated for each comparison. Sliding window profiles of COI-COII nucleotide divergence (bold) and percent transitions (fine) are presented for each species pair. Left axes indicate percent nucleotide divergence, and right axes represent percent transitions. Areas of < 50% transitions (dotted line) are shaded to indicate regions of saturation. Arrows indicate regions of relatively high divergence and high percent transitions. **85**

Figure 4- 1: Locations of specimens examined. Filled circles indicate specimens used for both molecular (DNA) and morphological (morph) surveys, and open circles indicate localities of specimens surveyed for morphological characters alone. Three locations not shown are sites in Anchorage AK (DNA+morph), Timmins ON (morph), and Germany (DNA+morph). The shaded area indicates the region of sympatric distribution, with dashed line separating the putative ranges of *Dioryctria reniculelloides* (northeastern distribution) and *D. pseudotsugella* (southwestern distribution). **118**

Figure 4- 2: Forewing measurements used for morphometric analysis of morphological variation. Length: Maximum forewing length, not including fringe. ASBA: Area of the pinkish-orange subbasal patch. Mode: Modal greyscale colour of pinkish-orange subbasal patch. ADS: Area of dark scales for the entire surface of the forewing, excluding fringe. **119**

Figure 4- 3: Strict consensus tree of 16 most parsimonious trees (length = 96) of 22 COI haplotypes representing 96 specimens from 22 localities in Alberta, British Columbia, Alaska and Oregon. Branch support is indicated by bootstrap values if over 50%. COI haplotype, number of specimens and collecting locality are indicated on terminal nodes. *D. auranticella* and *D. rossi* were used as outgroup taxa. **120**

Figure 4- 4: A nested parsimony network of 22 COI haplotypes representing 96 individuals. Hollow circles represent individual haplotypes, with size proportional to the number of specimens. Solid circles indicate missing haplotypes, based on the

parsimony analysis. Lines linking haplotypes represent single nucleotide changes. Dashed numbers indicate the nesting pattern of the haplotypes, with dashed boxes representing one-step clades, thin solid lines are two-step clades, and dark solid lines are three-step clades. Haplotype J was found to have the greatest root possibility, based on parsimony analysis. 121

Figure 4- 5: A semi-strong hybrid multidimensional ordination (SSH MDS) of four morphometric forewing characters: Length (maximum forewing length), ADS (area dark scales), ASBA (area of the pinkish-orange subbasal area), and Mode (modal grayscale colour value of scales in the subbasal area). Three extrinsic characters were also overlain as vectors (Sex, mtDNA group, and host plant). The ordination showed significant clustering when grouped by host plant and all vectors shown are also significant at $P<0.05$. 122

Chapter 1: General Introduction

According to Mayr (1969, pg. 2) systematics "includes not only the service functions of identifying and classifying but the comparative study of all aspects of organisms, as well as interpretation of the role of lower and higher taxa in the economy of nature and evolutionary history." This statement serves to highlight two important aspects of the field of systematics. First, systematics strives to interpret evolutionary relationships among organisms and examine the processes that have led to their diversity. Second, systematics provides the means to delimit, describe, and identify species in nature. Although this process of identifying and classifying organisms is viewed by many (as indicated above) as a service to other fields of biology, describing and delimiting species is a rigorous, hypothesis-driven science (Lipscomb et al. 2003, Seberg et al. 2003, Wheeler 2004, Prendini 2005). Identifying species boundaries and testing species limits requires an in-depth knowledge of population genetics, phenotypic variability, molecular ecology, and phylogeography. During the course of my thesis, I drew on all of these fields to clarify the systematics and improve identification of species in the moth genus *Dioryctria*.

The genus *Dioryctria*

First described by Zeller (1846), *Dioryctria* is a large, primarily Holarctic group, although several species have been described from the northern neotropics (Neunzig and Dow 1993, Neunzig 1996). Heinrich (1956) observed that "*Dioryctria* is one of the most, if not the most, distinct and sharply defined genus in the Phycitidae; and is so despite the variations in structure exhibited by its male antennae, male maxillary palpi, and its wing scaling." Members of the genus are easily recognized by their characteristic forewing pattern and genitalic structure (Fig. 1-1). The presence of a white discocellular spot and two pale dentate transverse bands on the forewing are particularly useful diagnostic features (Heinrich 1956, Neunzig 2003).

Currently, the genus is composed of 79 recognized species, although high levels of forewing pattern variation and a lack of diagnostic genitalic characters have hindered

species-level identification in this well-defined genus. To improve identification, 11 species groups have been recognized (Heinrich 1956, Mutuura and Munroe 1972, 1974, Wang and Sung 1982, Segerer and Prose 1997, Speidel and Asselbergs 2000, Neunzig 2003) that help to delineate morphological variation among species. The seven most speciose species groups (*abietella, auranticella, baumhoferi, ponderosae, schuetzeella, sylvestrella,* and *zimmermani* groups) have been supported by molecular and morphological characters (Du et al. 2005), although several species groups require increased taxon sampling and relationships among these groups remain unclear.

Erecting species groups has been a useful method to circumscribe the morphological variation within this large genus, but species identification within these species groups continues to be problematic (Sopow et al. 1996). Many species occur sympatrically, and are separated primarily by larval host plant association or minor morphological differences. However, overlapping morphological variation has been well documented, bringing into question the distinctness of many species (Sopow et al. 1996).

Dioryctria larvae specialize on a wide variety of coniferous hosts, and target regions of rapid growth, such as cones, cambial tissue, buds, shoots, and foliage (Neunzig 2003). Host association is often limited to one or two host plants, and is often diagnostic (Mutuura et al. 1969a, 1969b, Mutuura and Munroe 1973, Neunzig 2003). Feeding damage by *Dioryctria* species can result in significant economic losses to the forestry industry (Lyons 1957, Hedlin et al. 1980, Hainze and Benjamin 1984, Blake et al. 1989, Mosseler et al. 1992, Turgeon et al. 1994) and has increased the need for accurate species diagnosis to support monitoring and control (Grant et al. 1987, Grant et al. 1993, Millar et al. 2005).

Diagnostic Characters

In groups such as *Dioryctria*, a number of characters are often needed to improve species diagnostics and delimitation. A number of characters have proven useful for improving species delineation and diagnostics, and I will discuss the major character sources used throughout this thesis, and comment on their advantages and disadvantages. In order to assess the success and accuracy of diagnostic character systems, it is

important to clarify what species definition is being applied. Throughout this thesis I consider a species to be a cohesive group of populations that maintains its genomic integrity inspite of gene flow with other groups, and where allopatric populations may also be considered separate species if they exhibt levels of character divergence similar to distinct, sympartic species (Sperling 2003).

Molecular Markers

Mitochondrial DNA (mtDNA) is used extensively throughout this thesis to improve diagnostics (Chapters 2 & 4), explore patterns of molecular evolution within and between sister species (Chapter 3), and infer species boundaries (Chapter 4). The advantages of using mtDNA genes are detailed in the subsequent chapters, but briefly are: maternal uniparental inheritance, lack of recombination, rapid evolution and robustness to degradation (Avise et al. 1987). There are, however, several disadvantages to relying on mtDNA loci for species-level examinations (Shaw 2002, Zhang and Hewitt 2003, Ballard and Whitlock 2004). First and foremost, mtDNA is a single gene complex and provides a single genealogy. Gene trees inferred from mtDNA characters can have a significantly different evolutionary history from the actual species tree, thus caution must be used when relying solely on mtDNA to infer relationships among taxa. Second, mtDNA may be susceptible to introgression due to interspecific hybridization (see Chapter 5). Third, many mtDNA genes often have heavily biased base composition (Ballard and Whitlock 2004). Fourth, nuclear encoded mtDNA pseudogenes are known to exist and can lead to incorrect gene trees if not detected (Zhang and Hewitt 1996).

Nuclear genes have been used much less extensively than mtDNA for delimiting species, but can be an excellent independent data source for inferring species boundaries (Zhang and Hewitt 2003). Nuclear loci suffer less from biased base composition and often have a more homogeneous substitution rate than mtDNA genes (Lin and Danforth 2003). Unfortunately, nuclear genes are more suceptible to degradation, frequently have allelic variants, evolve more slowly, and can be susceptible to gene duplication (Zhang and Hewitt 2003).

Given that each type of molecular marker has its own advantages and disadvantages, the true strength of molecular characters lies in their ability to provide independent assessment of species boundaries when used together in multi-locus studies. Incongruence caused by introgression or incomplete lineage sorting can be detected if independent loci are examined, a phenomenon that would otherwise be missed when loci are examined separately.

Morphological and Ecological Characters

Traditional taxonomists rely heavily on a combination of morphological, behavioural and ecological characters to delimit and diagnose species. Difficulties with these characters arise when species are separated by little or no variation (i.e. cryptic species), or are highly variable throughout their range (i.e. polymorphic species or host races) (see Chapters 2 and 4). The loss of taxonomic expertise and the time required to identify and describe unknown species using these types of characters is significant, and has become known as the taxonomic crisis (Godfray 2002). Even with these concerns, the morphological, behavioural and ecological differences among species are an essential component of accurate species delimitations (Ebach and Holdrege 2005a, Prendini 2005). These characters are needed to diagnose species in the field, and are the only way to link specimens to previously described species (see Chapter 2), or to reliably connect names to other knowledge about an organism (Ebach and Holdrege 2005b). These types of characters are particularly good for testing species boundaries inferred by molecular markers, and visa versa. Ultimately, an integrated approach to species delimitations is needed to obtain the broadest, most applicable survey of species boundaries (Wheeler 2004), and the need for combined data is emphasized throughout this thesis.

Scope of the thesis

Ultimately, my goal for this thesis is to explore several important facets of species identification, diagnosis and delimitation in the genus *Dioryctria.* In Chapter 2, I examine the correlation of mtDNA variation and phylogenetic relationships to

morphological, behavioural and ecological variation, and use a holistic approach to identify useful diagnostic characters for a suite of *Dioryctria* species in a seed orchard in northern California. In Chapter 3, I explore patterns of mtDNA nucleotide variability within and among species of Lepidoptera and Diptera and will discuss the implications to mtDNA-based methods of species delineation. In Chapter 4, I test the traditional delimitation of two broadly sympatric *Dioryctria* species using three independent molecular markers, morphometric characters, geographic distribution and larval host plant associations. I determine species boundaries using estimations of gene flow and genetic diversity among populations of these *Dioryctria* throughout western North America.

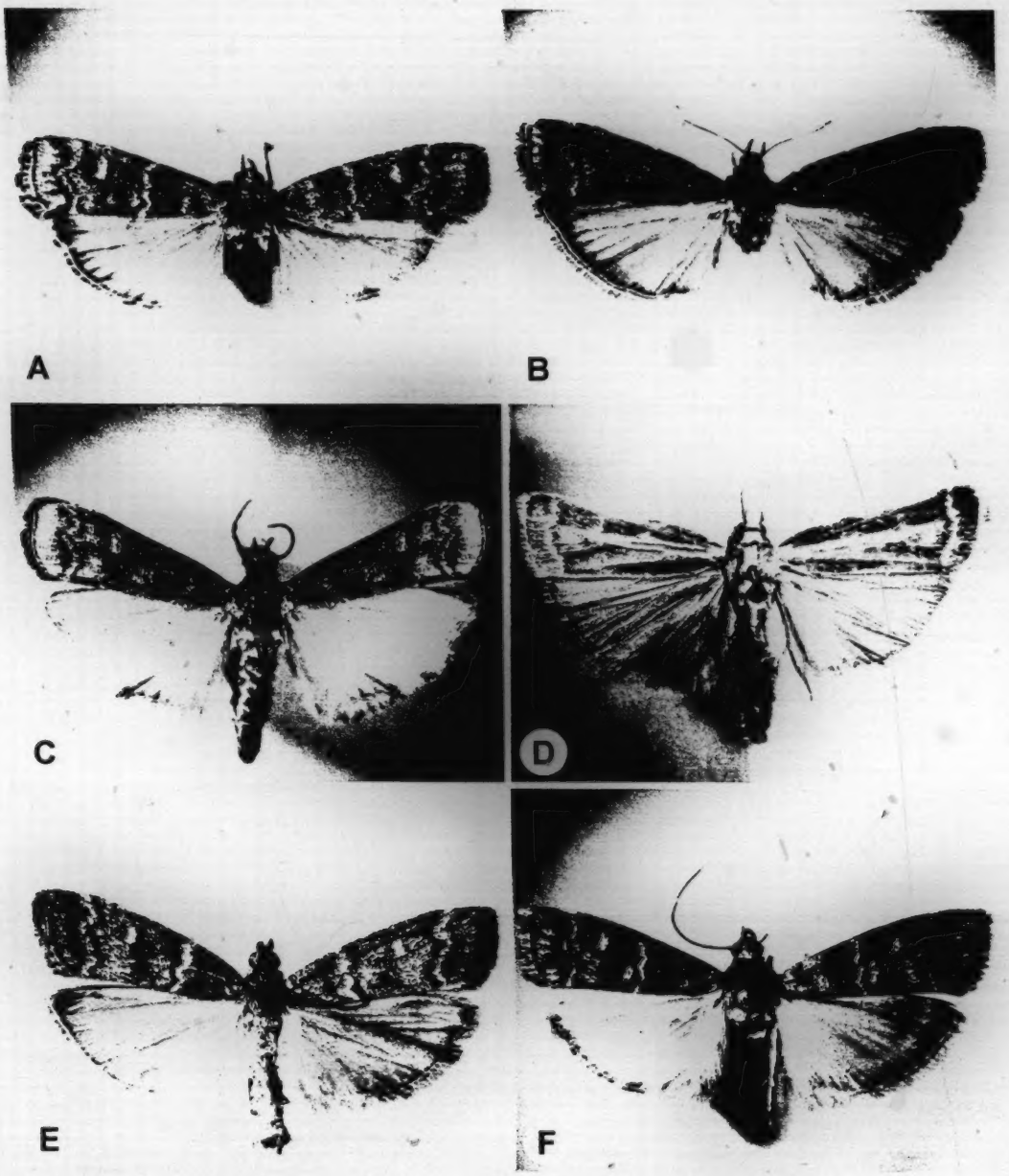

Figure 1-1: Common western North American *Dioryctria* species.
A) *D. pentictonella*; B) *D. cambiicola*; C) *D. fordi*; D) *D. rossi*; E) *D. pseudotsugella*; F) *D. okanaganella.*

Literature Cited

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18: 489-522.

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729-744.

Blake, E. A., M. R. Wagner, and T. W. Koeber. 1989. Relative effect of seed and cone insects on ponderosa pine in northern Arizona. Journal of Economic Entomology 82: 1691-1694.

Du, Y., A. D. Roe, and F. A. H. Sperling. 2005. Phylogenetic framework for *Dioryctria* (Lepidoptera: Pyralidae: Phycitinae) based on combined analysis of mitochondrial DNA and morphology. The Canadian Entomologist 137: 685-711.

Ebach, M. C., and C. Holdrege. 2005a. DNA barcoding is no substitute for taxonomy. Nature 434: 697.

Ebach, M. C., and C. Holdrege. 2005b. More taxonomy, not DNA barcoding. BioScience 55: 822-823.

Godfray, H. C. J. 2002. Challenges for taxonomy. Nature 417: 17-19.

Grant, G. G., Y. H. Prevost, K. N. Slessor, G. G. King, and R. J. West. 1987. Identification of the sex pheromone of the spruce coneworm, *Dioryctria reniculelloides* (Lepidoptera: Pyralidae). Environmental Entomology 16: 905-909.

Grant, G. G., S. A. Katovich, D. J. Hall, D. A. Lombardo, and K. N. Slessor. 1993. Sex-pheromone identification and trapping of *Dioryctria resinosella* (Lepidoptera, Pyralidae). Environmental Entomology 22: 154 -161.

Hainze, J. H., and D. M. Benjamin. 1984. Impact of the red pine shoot moth, *Dioryctria resinosella* (Lepidoptera: Pyralidae), on height and radial growth in Wisconsin red pine plantations. Journal of Economic Entomology 77: 36-42.

Hedlin, A. F., H. O. Yates III, D. C. Tovar, B. H. Ebel, T. W. Koerber, and E. P. Merkel. 1980. Cone and seed insects of North American conifers. Canadian Forestry Service.

Heinrich, C. 1956. American Moths of the Subfamily Phycitinae. Smithsonian Institution, Washington D.C.

Lin, C.-P., and B. N. Danforth. 2003. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Molecular Phylogenetics and Evolution 30: 686-702.

Lipscomb, D., N. Platnick, and Q. Wheeler. 2003. The intellectual content of taxonomy: a comment on DNA taxonomy. Trends in Ecology and Evolution 18: 65-66.

Lyons, L. A. 1957. Insects affecting seed production in red pines. II. *Dioryctria disclusa* Heinrich, *D. abietella* (D and S) and *D. cambiicola* (Dyar) (Lepidoptera: Phycitidae). The Canadian Entomologist 89: 70-79.

Mayr, E. 1969. Principles of Systematic Zoology. McGraw-Hill, New York, NY.

Millar, J. G., G. G. Grant, J. S. McElfresh, W. Strong, C. Rudolph, J. D. Stein, and J. A. Moreira. 2005. (3Z, 6Z, 9Z, 12Z, 15Z)-pentacosapentaene, a key pheromone component of the fir coneworm moth, *Dioryctria abietivorella.* Journal of Chemical Ecology 31: 1229-1234.

Mosseler, A., B. A. Roberts, and P. Tricco. 1992. The effects of fir coneworm, *Dioryctria abietivorella* (Grote) (Lepidoptera: Pyralidae), on seed production in small isolated populations of red pine, *Pinus resinosa* Ait. Forest Ecology and Management 53: 15-27.

Mutuura, A., and E. Munroe. 1972. American species of *Dioryctria* (Lepidoptera: Pyralidae) III. Grouping of species: Species of the *auranticella* group, including the Asian species, with the description of a new species. The Canadian Entomologist 104: 609-625.

Mutuura, A., and E. Munroe. 1973. American species of *Dioryctria* (Lepidoptera: Pyralidae) IV. The *schuetzeella* group and the taxonomic status of the spruce cone moth. The Canadian Entomologist 105: 653-6[illegible]

Mutuura, A., and E. Munroe. 1974. A new genus related to *Dioryctria* Zeller (Lepidoptera: Pyralidae: Phycitinae), with definition of an additional species group in *Dioryctria.* The Canadian Entomologist 106: 937-940.

Mutuura, A., E. Munroe, and D. A. Ross. 1969a. American species of *Dioryctria* (Lepidoptera: Pyralidae) I. Western Canadian species of the *zimmermani* Group. The Canadian Entomologist 101: 1009-1023.

Mutuura, A., E. Munroe, and D. A. Ross. 1969b. American species of *Dioryctria* (Lepidoptera: Pyralidae) II. Western Canadian species of the *baumhoferi* and *ponderosae* groups. The Canadian Entomologist 101: 1042-1047.

Neunzig, H. H. 1996. New species of Phycitinae (Lepidoptera: Pyralidae) from the Dominican Republic. Proceedings of the Entomological Society of Washington 98: 774-801.

Neunzig, H. H. 2003. Pyraloidea, Pyralidae (part), Phycitinae (part). The Moths of America North of Mexico fasc. 15.5.

Neunzig, H. H., and L. C. Dow. 1993. The Phycitinae of Belize (Lepidoptera: Pyralidae). North Carolina Agricultural Research Service, North Carolina State University, Raleigh, NC.

Prendini, L. 2005. Comment on "Identifying spiders through DNA barcodes". Canadian Journal of Zoology 83: 298-304.

Seberg, O., C. J. Humphries, S. Knapp, D. W. Stevenson, G. Petersen, N. Scharff, and N. M. Andersen. 2003. Shortcuts in systematics? A commentary on DNA-based taxonomy. Trends in Ecology and Evolution 18: 63-65.

Segerer, A. H., and H. Prose. 1997. *Dioryctria resiniphila* sp. n., eine neue Pyralidae auf *Abies cephalonica* LOUD. in Griechenland. NachrBl. bayer. Ent. 46: 57-67.

Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proceedings of the National Academy of Sciences of the United States of America 99: 16122-16127.

Sopow, S. L., R. G. Bennet, J.-F. Landry, and B. Landry. 1996. Identification of the 'grey' *Dioryctria* species of British Columbia (Lepidoptera: Pyralidae). Journal of the Entomological Society of British Columbia 93: 75-91.

Speidel, W., and J. E. F. Asselbergs. 2000. The status of *Ocrisia* Ragonot, 1893, and notes on *Dioryctria* Zeller, 1846 (Lepidoptera: Pyralidae, Phycitinae). Entomologische Zeitschrift 110: 144-146.

Sperling, F. 2003. Butterfly molecular systematics: from species definitions to higher-level phylogenies, pp. 431-458. *In* C. Boggs, W. Watt and P. Ehrlrich [eds.], Butterflies: Evolution and Ecology Taking Flight. University of Chicago, Chicago Press, Chicago.

Turgeon, J. J., A. Roques, and P. de Groot. 1994. Insect fauna of coniferous seed cones: diversity, host plant interactions, and management. Annual Review of Entomology 39: 179-212.

Wang, P.-Y., and S.-M. Sung. 1982. Description of a new species of *Dioryctria* Zeller on *Pinus sylvestris* var. Mongolica from north-east China, with establishment of a new species group. Acta Entomologica Sinica 25: 324-327.

Wheeler, Q. D. 2004. Taxonomic triage and the poverty of phylogeny. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 359: 571-583.

Zeller, P. C. 1846. Die Knotenhornigen Phyciden. Isis von Oken 39: 729-788.

Zhang, D.-X., and G. M. Hewitt. 1996. Nuclear integrations: Challenges for mitochondrial DNA markers. Trends in Ecology and Evolution 11: 247-251.

Zhang, D.-X., and G. M. Hewitt. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular Ecology 12: 563-584.

Chapter 2: Identification of *Dioryctria* (Lepidoptera: Pyralidae) in a seed orchard at Chico, California

*A version of this chapter is published.

Roe, A.D.R., J.D. Stein, N.E. Gillett, and F.A.H. Sperling. 2006. Annals of the Entomological Society of America. 99(3): 433-448.

Introduction

Dioryctria Zeller (Lepidoptera: Pyralidae) is a Holarctic genus currently composed of 70 species, with 40 species described from North America north of Mexico (Heinrich 1956, Munroe 1959, Mutuura et al. 1969b, 1969a, Schaber and Wood 1971, Coulson et al. 1972, Mutuura and Munroe 1972, 1973, 1979, Mutuura 1982, Blanchard and Knudson 1983, Neunzig and Leidy 1989) and at least six additional species from Mexico (Cibrián-Tovar et al. 1986, Neunzig 1990). The distribution of the genus matches that of the coniferous hosts used by members of the species, which range from subtropical to subarctic coniferous forests (Neunzig 2003).

Larvae of the majority of species feed internally on coniferous trees, attacking regions with rapid growth (meristematic tissue) such as cones, stems, cambium, wounds and blister rust galls (e.g., *Endocronartium harkenssii* (Moore) Y. Hiratsuka and *Cronartium coleosporioides* Arthur), although several species feed externally on foliage and buds (Neunzig 2003). Damage can lead to substantial economic loss, particularly in tree farms, plantations, shelterbelts and seed orchards (Lyons 1957, Hedlin et al. 1980, Blake et al. 1989, Mosseler et al. 1992). Cone feeding destroys seeds, reducing seed production, especially during years of low seed set (Schowalter et al. 1985). Stem, trunk and shoot mining results in branch breakage and bud loss, causing tree deformation during heavy infestations (Hainze and Benjamin 1984).

Dioryctria infestations at the USDA Forest Service Genetic Resource Center (GRC) in Chico, California illustrate clearly the importance of these pests. The GRC supports 122.8 acres of grafted breeding stock and produces seedlings for three breeding zones of ponderosa pine (*Pinus ponderosa* Douglas ex Lawson & C. Lawson), two

breeding zones of sugar pine (*Pinus lambertiana* Douglas) and five breeding zones of Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco) that support reforestation efforts throughout northern California. Twenty years of heavy *Dioryctria* infestations have drastically reduced the number of viable seeds harvested from this orchard, hindering its ability to produce seedlings (G. Norcross, personal communication). Traditional methods of managing insect infestations, such as insecticide sprays and injections, have resulted in inconsistent control, so alternative methods such as pheromone monitoring and mating disruption are being pursued.

Development of pheromone monitoring and control methods requires accurate species identification. As a genus, adult *Dioryctria* are one of the most easily identified groups in the subfamily Phycitinae (Heinrich 1956). Seven species groups delineated by genitalic and forewing characters were originally erected to improve the taxonomy and aid identification of this difficult group (Mutuura and Munroe 1972). Two additional North American species groups have been delineated (Neunzig 2003); however, identification and taxonomy of many *Dioryctria* species within and between groups remains problematic. Previous taxonomic work in North America has relied upon minor genitalic variation, slight differences in forewing pattern, geographic range and larval host associations. Wing pattern differences can be polymorphic, compounding the problem. Also, many species are sympatric, and occur on the same larval host. Thus it is difficult to rely on these characters for species identification (Sopow et al. 1996), and additional characters are needed to reliably identify *Dioryctria* species.

Nucleotide sequence data, particularly mitochondrial DNA (mtDNA), have been useful in resolving difficult species problems by providing a suite of additional characters (Simon et al. 1994, Caterino et al. 2000). MtDNA is maternally inherited, and essentially haploid. Mutations can accumulate rapidly, allowing the separation of closely related species. MtDNA is also robust to degradation, permitting the use of pheromone-trapped material and some museum specimens. The cytochrome-c oxidase gene regions I and II (COI, COII) have been shown to be particularly useful for resolving species problems across a range of lepidopteran families (Caterino et al. 2000, Sperling 2003). The use of COI for identifying species has been popularized by DNA barcoding (Hebert et al. 2003), although the use of this technique as the sole method for identifying species has been

questioned (Lipscomb et al. 2003, Sperling 2003b, Will and Rubinoff 2004). Previous molecular work on *Dioryctria* has used isozymes, cuticular hydrocarbons, and nucleotide sequence data to examine genetic variation within and between *Dioryctria* species (Richmond 1995, Richmond and Page 1995, Knölke et al. 2005, Du et al. in press), but has not focused on the full complement of species likely to be encountered in any one region.

This study is intended to provide a foundation for molecular identification, and to clarify and confirm morphological, larval host association and pheromone characters that may be used for identification in the field. There were four primary objectives of this study. First, I used a 475 bp region of COI to identify distinct mtDNA lineages of *Dioryctria* from the Genetic Resource Center in Chico, CA, including additional specimens from northern California, Oregon and British Columbia. Second, I associated these genetic lineages with previously described species using adult morphology, locality, larval host association and pheromone attraction. Third, I sequenced 2.3 kb from each major genetic lineage and obtained a well supported preliminary phylogeny for species and species groups identified in the region. Fourth, I used subsets of the 2.3 kb sequences to compare divergences between the 475 bp fragment used in this study and the 658 bp DNA barcoding region of Hebert et al. (2003) to examine the information content of these regions in *Dioryctria* species. The overall aim of this project was to use mtDNA sequence data to identify *Dioryctria* species, evaluate boundaries and provide a preliminary assessment of the phylogenetic relationships in the genus.

Materials and Methods

Collection Sites and Species

A total of 180 *Dioryctria* specimens were collected from northern California, Oregon and British Columbia (Table 2-1). Collecting was focused on the US Forest service conifer seed orchard at GRC in Chico, CA, and 146 specimens were collected at this locality. An additional 34 specimens collected from northern California, western Oregon, and British Columbia were included to expand species sampling and provide an assessment of geographic variation. Samples were collected to represent the range of

pheromone attraction, larval host plant association, and morphology that was considered likely to be seen in *Dioryctria* species throughout the region, especially from northern California and Oregon.

All six species groups previously recorded in northern California were represented in the study. Eight species of *Dioryctria* were sampled in this survey, with four recorded at Chico and six in northern California (Table 2-1): *D. abietivorella* (Grote) (*abietella* group); *D. auranticella* (Grote) and *D. rossi* Munroe (*auranticella* group); *D. pentictonella* Mutuura, Munroe & Ross (*baumhoferi* group); *D. okanaganella* Mutuura, Munroe & Ross (*ponderosae* group); *D. pseudotsugella* Munroe (*schuetzeella* group); *D. cambiicola* (Dyar) and *D. fordi* Donahue and Neunzig (*zimmermani* group). *Dioryctria* specimens were identified based on wing characters, genitalic morphology, and geographic range using keys and other published materials (Heinrich 1956, Munroe 1959, Mutuura et al. 1969a, 1969b; Mutuura and Munroe 1972, 1973; Neunzig 2003, Donahue and Neunzig 2005). Four additional species of *Dioryctria* have been recorded from northern California (Neunzig 2003), but were not included because they are generally rare in the region and no fresh material could be obtained for this study. *Dioryctria muricativorella* Neunzig, *D. mutuurai* Neunzig, and *D. westerlandi* Donahue and Neunzig have been described recently from California, but have highly restricted geographic ranges and few specimens are known. *Dioryctria ponderosae* Dyar was not collected during this study, although this species is recorded at low frequency at various sites in the study region (Neunzig 2003). Two additional species in the Phycitini, *Oncocera faecella* (Zeller) and *Ceroprepes ophthalmicella* (Christoph) (Lepidoptera: Pyralidae: Phycitinae), were included as outgroup taxa.

Collection Methods

Specimens examined in this study were provided by collaborators or collected by the authors (Table 2-1). Larvae were extracted or reared from cones, cambial tissue and pitch masses from eight conifer species, with the majority of material obtained from the GRC. Live larvae and reared adults were preserved in 96-100% ethanol. Adults were sampled using both light and pheromone-baited traps. Live light-trapped specimens were frozen at –20 °C or –70°C or were placed in 96-100% ethanol.

Dioryctria specimens were collected by pheromone trapping at GRC in 1998, 2000, and 2001. These lures differed either in chemical composition or concentration. Pheromone lures were developed based on previously described lures for *Dioryctria disclusa* Heinrich (Meyer et al. 1982), *D. abietella* (Denis and Schiffermüller) (Löfstedt et al. 1983), *Dioryctria clarioralis* (Walker) (Meyer et al. 1984), *Dioryctria amatella* (Hulst) (Meyer et al. 1986), *Dioryctria reniculelloides* Mutuura and Munroe (Grant et al. 1987), and *Dioryctria resinosella* Mutuura (Grant et al. 1993). Pheromone trapping with undefined blends obtained a small number of specimens in 1998, but since these lures could not be confidently associated with chemical blends they were not considered in the pheromone analysis. Pherocon 1CP pheromone traps baited with the lures (Trécé Inc., Salinas, CA) were placed in a replicated, randomized block pattern in tree tops throughout the orchard to maximize trap catch (Grant et al. 1987). Traps were checked biweekly and specimens removed, scored for wing pattern and frozen at –20°C.

Molecular Techniques

Genomic DNA was extracted from thoracic muscle or legs of specimens using a QIAamp DNA Mini Kit (250) (Qiagen, Valencia, CA), and visualized on 0.8% agarose gels (Gibco BRL). MtDNA was amplified using a Polymerase Chain Reaction (PCR) with pairs of heterologous primers (Simon et al. 1994) on either a Whatman Biometra TGradient or TPersonal Thermocycler (Whatman Biometra, Göttingen, Germany) with Taq polymerase (University of Alberta, Edmonton, Alberta, Canada) added in a hot start at the end of an initial denaturation cycle at 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 45°C for 30 s, 72°C for 2 min, and a final extension at 72°C for 5 min. A 475 bp fragment in the COI region was obtained using either Jerry or Jerry V and Mila III (Table 2-2) for 180 specimens in the study region. From the specimens examined for 475 bp, 14 specimens representing eight distinct genetic lineages in the region were sequenced over the full 2.3 kb of COI-COII. A complete list of primers used to obtain the 2.3 kb fragment is shown in Table 2-2. PCR products were cleaned using QIAquick PCR Purification Kit (250) (Qiagen, Valencia, CA) and cycle sequenced on either a Whatman Biometra TGradient or TPersonal Thermocycler using Amersham Bioscience DYEnamic ET Dye Terminator Kit (Amersham Biosciences, Buckinghamshire, England)

according to the following profile: initial denaturation at 93.0°C for 30 s, 28 cycles of 95°C for 20 s, 45°C for 15 s, 60°C for 1 min, with a final extension at 60°C for 30 s. The sequenced product was purified by filtration through Sephadex columns (Amersham Biosciences, Uppsala, Sweden) and dried. This product was re-suspended in formamide and sequenced on an ABI 377 automated sequencer (Applied Biosystems, Foster City, CA). All fragments were sequenced in both directions and manually aligned to the sequence of *Drosophila yakuba* Burla (Clary and Wolstenholme 1985).

Phylogenetic analyses of DNA sequence

Initially the 475 bp COI fragments from all 180 individuals were compared and identical mtDNA haplotypes were combined. A total of 28 unique haplotypes of *Dioryctria* were found among all specimens sampled (Table 2-1). Sequences were aligned by eye and nucleotides were treated as unordered, unweighted characters. Phylogenetic analysis using maximum parsimony (MP) was conducted with PAUP 4.0*b10 (Altivec) (Swofford 2002). MP analysis was performed using heuristic searches with the following parameters: 100 random addition replicates; stepwise addition; tree bisection-reconnection (TBR). Branch support was calculated using bootstrap and Bremer support values. Bootstrap values were obtained with 100 bootstrap replicates using heuristic search methods as described above. Bremer support was calculated from a strict consensus MP tree using AutoDecay 5.03 (Eriksson 2002). Analysis of the 2.3 kb of sequence data for the 14 available specimens was identical to that described for the 475 bp sequence data.

Morphological Techniques

A wing phenotype classification was developed with the intent of linking forewing pattern to genetic variation, seasonality or pheromone lure attraction. Specimens were initially scored by C. Rudolf (USDA FS. PSW Research Station) as wing pattern types that were identified and described by J. D. Stein, in conjunction with additional USDA funded *Dioryctria* surveys conducted prior to this research project. Specimens were subsequently characterized using a simplified wing phenotype character system developed by A. D. Roe, in which forewing pattern variability was scored for

three characters: presence of raised scales, primary forewing colour and colour of subbasal area. Colours were standardized against the Munsell Soil Colour Charts 1994 revised edition (Munsell Colour, New Windsor, NY) and colour scores are listed after each colour description. Character 1, presence of raised scales, was scored as present (R) or absent (N). Raised scales, when present, occurred as a patch in the basal area, as a subbasal ridge, a medial ridge and on the discocellular spot. Character 2, primary forewing colour, was scored as black to dark grey (B) (2.5/N to 6/N), white to pale grey (W) (7/N to 8/N), orange (O) (7.5YR 7/8) or brown (Br) (10YR 6/2 to 5/2). Wings that appeared white or pale grey often had white tipped scales with dark interiors, lightening the overall appearance of the wing. Individuals with this colouration were scored (W) to reflect the pale appearance of the wing. Character 3, colour of subbasal area, was scored as black to dark grey (B) (2.5/N to 6/N), white to pale grey (W) (8/N to 7/N), or tan to dark red (R) (5YR 4/6 – 7/6 to 10R 3/6 – 5/6 to 2.5YR 3/6 – 6/6). Each specimen was then assigned a wing phenotype code that reflected all three characters. For example, a forewing with raised scales that is primarily black and has a red subbasal area would be RBR. Any larvae or specimens with damaged forewings were listed as unscorable (UNS). All specimens were scored by A.D. Roe and were confirmed by an independent observer.

Genitalia dissection methods were adapted from techniques outlined by Sopow et al. (1996) and Winter (2000). The abdomen was removed and placed in approximately 5 ml of 10% KOH solution and boiled in a beaker of water for 10 min or until the abdomen was softened. The abdomen was placed in a 30% ethanol solution and scales were removed using a No. 0 insect pin and a soft hair paintbrush. The genitalia were extracted using a pair of fine forceps and a No. 0 insect pin and disconnected from adjoining membranes. A glass capillary tube pulled to 5 μm in diameter was attached to microtubing and a fine syringe and was used to evert male vesicae. This apparatus was filled with 30% ethanol solution and then slowly injected into the aedeagus causing the vesicae to evert. Vesicae contained many cornuti that hindered the process, so a hooked No. 0 insect pin was used to help eversion. After examination, genitalia were placed in glycerin in genitalia microvials and pinned with the specimen voucher.

Images of wings and genitalia were taken with a Nikon COOLPIX 990 Digital Camera mounted on a dissecting microscope. Multiple images were taken of each specimen and compiled in AutoMontage (Syncroscopy, Frederick, MD). Wings, head capsule and remaining structures were preserved in gelatin capsules for morphological comparisons and future identifications. Vouchers and images are deposited in the E.H. Strickland Entomological Museum, University of Alberta, Edmonton, Alberta, Canada.

Results

Analysis of 475 bp COI Fragment

In the 475 bp fragment of COI examined in 180 specimens, 101 positions were variable and 77 were parsimony informative, with an AT bias of 68.8% and a transition:transversion ratio of 3.23. Six of the 28 haplotypes were found in at least two populations throughout the study range, some of which were separated by large geographic distances (Table 2-1). There were 22 haplotypes restricted to single localities, and 19 of these haplotypes were unique and found in a single individual.

A heuristic maximum parsimony search found 178 most parsimonious trees that were 182 steps in length. Multiple most-parsimonious trees resulted from rearrangements between similar haplotypes and between lineages. Eight major genetic lineages were obtained from specimens in the study region. Each of the eight lineages contained more than one specimen and were identified based on relatively long basal branch lengths, high bootstrap values (93 – 100%), and Bremer support values (3 – 14) (Figs. 2-1A and B). Some lineages (Lineages 1 and 8) contained 2 – 3 specimens, while Lineages 6 and 7 each contained over 70 specimens. (Table 2-1; Fig. 2-1).

Sequence divergences within and between lineages for the 475 bp sequences were measured using uncorrected pairwise distances. Divergences within lineages ranged from 0.0% to 1.7%. Divergences between lineages ranged from 3.6% (Lineages 1 to 2) to 8.0% (Lineages 2 to 6). Divergences between species groups ranged from 4.0% (*baumhoferi* Gr to *auranticella* Gr.) to 8.0% (*abietella* Gr. to *zimmermani* Gr.). Divergences between outgroup taxa and the ingroup ranged from 6.7% (RS1g to *C.*

ophthalmicella; RS1g to *O. faecella*; RS1b to *O. faecella*) to 9.9% (RE5 to *C. ophthalmicella*; AB2 to *C. ophthalmicella*).

Support for relationships between lineages was variable. Lineages 1 + 2 formed a well supported sister group relationship, but relationships between the remaining lineages were poorly resolved (Fig. 2-1B). Monophyly of *Dioryctria* was not supported by any of the most-parsimonious trees. The outgroup taxon *C. ophthalmicella* grouped with Lineage 7 in all trees, although this relationship was not supported by bootstrap or Bremer support values.

Morphological Identifications.

Forewings of all voucher specimens used for mtDNA sequencing were examined by A. Roe for structural characters (raised scales), colour, and wing pattern. A total of 146 specimens were scored for wing phenotype, while the remaining 34 individuals were unscorable (UNS), due to damage from sticky traps or because specimens were collected as larvae. Forewings were assigned a three-letter phenotype code based on the presence or absence of raised scales, primary forewing colour, and colour of the subbasal area (Table 2-3). Specimens initially formed two distinct groups: specimens lacking raised scales (N) and those with raised scales (R). For specimens lacking raised scales, three main groups of primary forewing colour were present: black, orange, or brown. These three groups corresponded to Lineage 6 (mainly NBB), Lineage 5 (NOR), Lineage 8 (NOR) and Lineage 4 (NBrR). For specimens with raised scales, forewings were either black or white. Wing phenotype codes could not distinguish between the raised scale lineages (Lineages 1, 2, 3, and 7) due to variability in forewing pattern in Lineage 7 (*D. pentictonella*).

Genitalic morphology and additional wing characters linked the eight lineages to eight previously described species from six species groups described from North America (Mutuura and Munroe 1972).

Lineage 1 (3 specimens) was identified as *D. cambiicola* (*zimmermani* group). These individuals could not be fully scored for wing phenotype due to forewing damage.

Lineage 2 (4 specimens) was identified as *D. fordi* (*zimmermani* group). Specimens in this lineage all had a wing phenotype code RWR, indicating the presence of

raised scales (R) with a primarily white forewing (W) and red subbasal area (R) (Table 2-3).

Lineage 3 (4 specimens) was identified as *D. okanaganella* (*ponderosae* group). Three specimens were scored for wing phenotype. Each specimen had raised scales (R) and a primarily black forewing (B), but variation in the colour of the subbasal area (black (B) or red (R)) produced two phenotypes for this species (RBB and RBR) (Table 2-3). The phenotypes scored for this species overlapped with those scored for Lineage 7.

Lineage 4 (6 specimens) was identified as *D. pseudotsugella* (*schuetzeella* group). The identification of *D. pseudotsugella* was based primarily on geographic range, rather than morphology, due to overlap of diagnostic characters with *D. reniculelloides*, a morphologically similar species. Although they are similar, mtDNA sequence data suggest that two sister taxa are distinct species (A.D.R., unpublished data). Five specimens were scored for wing phenotype. All scored specimens had a phenotype code NBrR, indicating the absence of raised scales (N), a primarily brown forewing (Br) and a tan to orange subbasal area (R) (Table 2-3).

Lineage 5 (7 specimens) was identified as *D. auranticella* (*auranticella* group). Five specimens were scored for wing phenotype. All scored specimens had a phenotype code NOR, indicating the absence of raised scales (N), a primarily orange forewing (O) and a reddish-orange subbasal area (R). The wing phenotypes scored for this lineage are identical to those scored for Lineage 8.

Lineage 6 (76 specimens) was identified as *D. abietivorella* (*abietella* group) and the majority were reared from hosts at the GRC. A total of 51 specimens were scored for wing phenotype. The majority of scored specimens had phenotype NBB, indicating the absence of raised scales (N), a primarily black forewing (B) and a black subbasal area (B). Nine additional specimens had a phenotype code NBR, indicating the presence of a tan to pale yellow patch in the subbasal area. Three specimens had a pale white subbasal area (NBW) and a single specimen had a primarily pale grey forewing (NWB). Specimens could be separated from all other species at the GRC based on the previously described wing phenotypes.

Lineage 7 (78 specimens) was identified as *D. pentictonella* (*baumhoferi* group). The majority of specimens were males collected at pheromone traps at the GRC in Chico,

CA, and equal numbers of males and females were reared or collected at lights. A total of 73 specimens were scored for wing phenotype. There was considerable overlap between the wing phenotypes found in *D. pentictonella* and those in other raised scale lineages. Six phenotypes were found in specimens of *D. pentictonella,* with RBB, RBR, and RWR the most common. All specimens had raised scales, but the primary colour of the forewing ranged from solid black to nearly white and the colour of the subbasal area was black, white or red. Based on collection dates for all adult material in the study, *D. pentictonella* appears to exhibit three peak flight periods (April 16 to June 16; July 1 to Aug 16; Sept 1 to Oct 16) (Fig. 2-2), although increased sampling is needed to confirm this phenology. *Dioryctria pentictonella* wing phenotypes were grouped by these three periods to determine if morphological variation was seasonal (Fig. 2-3). The three most common phenotypes (RBB, RBR, and RWR) were present in all three flight periods, while two phenotypes (RWW and RWB) were present in two periods. A single specimen with phenotype RBW occurred during a single flight period.

Lineage 8 (2 specimens) was identified as *D. rossi* (*auranticella* group) and contained two specimens. All specimens had a phenotype code NOR, indicating the absence of raised scales (N), a primarily orange forewing (O) and a reddish-orange subbasal area (R). The wing phenotypes scored for this lineage are identical to those scored for Lineage 5.

Larval Host Plant Associations

A total of 57 specimens in this study were reared or extracted as larvae from host plant material. Specimens were reared from cones, cambium or blister rust tissue on the following conifers: Douglas-fir (*Psuedotsuga menziesii)*, ponderosa pine (*Pinus ponderosa*), sugar pine (*Pinus lambertiana*), lodgepole pine (*Pinus contorta*), whitebark pine (*Pinus albicaulis* Engelmann), western white pine (*Pinus monticola*), Afghan pine (*Pinus brutia* ssp. *eldarica*) and western larch (*Larix occidentalis*).

There were 33 individuals reared from Douglas-fir cones and every specimen was identified as *D. abietivorella* (Lineage 6; Table 2-1). A total of 21 specimens were reared from cones of other conifer species, 18 of which were also identified as *D. abietivorella.* One specimen reared from an Afghan pine cone and two specimens reared from

whitebark pine cones were identified as *D. pentictonella* (Lineage 7). Three specimens were reared from cambial tissue or blister rust tissue in Douglas-fir and lodgepole pine and were identified as *D. cambiicola* (Lineage 1).

Pheromone Analysis

Specimens were trapped by five pheromone blends: I, II, III, IV, and V in 2000 and 2001 (Table 2-4). A total of 73 pheromone trapped specimens were sequenced, and represented the range of flight period, lure attraction and morphological variation occurring at the GRC. Three species of *Dioryctria* were trapped at pheromone lures: *D. pentictonella, D. abietivorella*, and *D. auranticella* (Table 2-4). The numbers of individuals captured varied substantially by lure (Table 2-4; Fig. 2-4). Lure I trapped a total of nine specimens from June 1st to Aug 31st. All nine individuals trapped by Lure I were identified as *D. abietivorella*, similar to the material reared from cones. Lure II trapped one specimen of *D. auranticella* and one of *D. abietivorella*. Lure III trapped seven specimens throughout the season and all were identified as *D. pentictonella*. Lure IV had the same chemical composition as Lure III, but with a lower dosage (Table 2-4), and trapped a single specimen of *D. pentictonella*. Lure V had the same chemical composition as the previous two lures, but with a higher dosage, and trapped 55 individuals, although the number of specimens trapped throughout the season varied (Fig. 2-4). All but a single specimen was identified as *D. pentictonella*. The other specimen caught by lure V was identified as *D. abietivorella*.

Phylogenetic Analysis of 2.3 kb fragment

Based on the eight genetic lineages recovered from the 475 bp fragment, 14 representative specimens were sequenced across the full length of COI and COII genes. Of the 2307 bp examined, 418 were variable and 309 were parsimony-informative, with an AT bias of 71.3% and a transition:transversion ratio of 2.66.

A heuristic maximum parsimony search found a single most-parsimonious tree 840 steps in length. A phylogram with bootstrap and Bremer support values is shown (Fig. 2-5). Eight distinct lineages were resolved by the 2.3 kb sequences, like the 475 bp

sequences. Single specimens were sequenced for most lineages, but where multiple specimens were used (Lineages 3, 5, 6 and 7) the lineages were well supported.

Sequence divergences among the 2.3 kb sequences were measured using uncorrected-pairwise distances. Divergences within lineages with more than one specimen ranged from 0.0% to 0.44%. Sequence divergence within species groups ranged from 2.2% (*zimmermani* Gr.) to 5.4% (*auranticella* Gr.). Divergences between species groups were also variable, ranging from 4.5% (*ponderosae* Gr. to *zimmermani* Gr.) to 7.5% (*abietivorella* Gr. to *ponderosae* Gr.). Divergence between outgroup and ingroup taxa ranged from 6.7% (*D. auranticella* to *O. faecella*) to 9.9% (*D. okanaganella* to *C. ophthalmicella*).

With the 2.3 kb sequences, species group and higher-level relationships were more resolved and showed increased support, compared to the 475 bp sequences (Figs. 2-1 and 2- 5). Well supported nodes (100% bootstrap values) from the 2.3 kb tree included the *zimmermani* Gr. (Node G), *ponderosae* Gr. (Lineage 3), *abietella* Gr. (Lineage 6) and *baumhoferi* Gr. (Lineage 7). The *auranticella* Gr. was paraphyletic, with *D. rossi* grouping with *D. pseudotsugella*, although this relationship was poorly supported (Node D). *Dioryctria auranticella* + (*D. rossi* + *D. pseudotsugella*) formed a poorly supported clade (Node C). The *abietella* Gr. (*D. abietivorella*) was sister to the remaining species groups, although this relationship was poorly supported (Node B). The *baumhoferi* Gr. was sister group to the *zimmermani* Gr. + *ponderosae* Gr. clade and this relationship was moderately supported (Node E). The sister group relationship between *zimmermani* Gr. and *ponderosae* Gr. was well supported (Node F). Monophyly of *Dioryctria* was resolved and well supported (Node A).

Information Content of Sequence Fragments

For specimens with the full 2.3 kb of COI+COII, uncorrected pairwise sequence divergences were compared between the 475 bp fragment used in this study and the 658 bp DNA barcoding region of Hebert et al. (2003). Divergences based on the full 2.3 kb of COI+COII were used as a reference. A wide degree of variability in sequence divergence between and within species groups was apparent between the two fragments (Fig. 2-6). Sequence divergences between species groups were generally higher in the

DNA barcoding fragment and lower in the 475 bp fragment than the full 2.3 kb sequence, although exceptions did occur. In two cases, the 475 bp fragment had slightly higher sequence divergence than both the barcoding and 2.3 kb fragments. Additionally, divergences as high as 1.7% were found in the larger data set of 475 bp sequences (RS1e to RS1d). Haplotype RS1e of *D. pentictonella* was collected at the same locality as many other *D. pentictonella* specimens, but it occurred later in the season than most other haplotypes (Table 2-1). Although sequence divergence was relatively high, haplotype RS1e was not considered a distinct lineage since it was only found in a single specimen and the haplotype was located terminally among other RS1 haplotypes (Fig. 2-1). Tree topologies of the two subsets (data not shown) were compared to the 2.3 kb tree (Fig. 2-5). Basal relationships in the 475 bp and DNA barcoding subsets were both poorly supported and lacked resolution. However trees from both subsets resolved all the well supported nodes found previously with the full set of 475 bp sequences and with the 14 sequences of 2.3 kb.

Discussion

MtDNA Lineages

DNA sequence from the 475 bp fragment of COI provided enough phylogenetic information to successfully delineate eight *Dioryctria* species in six species groups. By using a short fragment, I was able to survey large numbers of individuals and examine sequence variation across a range of morphological characters, pheromone lure associations and geographic locations. Previous phylogenetic studies among Lepidoptera have demonstrated the effectiveness of using mtDNA lineages for identifying distinct species (Landry et al. 1999, Kruse and Sperling 2001) and delineating species boundaries (Sperling et al. 1996, Landry et al. 1999, Sperling et al. 1999, Caterino et al. 2000, Kruse and Sperling 2001).

Preliminary tests demonstrated that known *Dioryctria* species were correctly delineated using the 475 bp fragment, so it was used to survey all specimens collected from the study region. This same region has also been used to investigate species problems in *Choristoneura* tortricids (Sperling and Hickey 1994). However, a different

658 bp region at the start of COI, known as the "DNA barcoding" region (Hebert et al. 2003) has more recently been used with increased frequency to identify closely related species or to associate specimens such as different sexes or immatures (Paquin and Hedin 2004, Simmons and Scheffer 2004, Hebert et al. 2005). Since the full COI gene was sequenced in a subset of *Dioryctria* specimens in order to improve resolution for the phylogeny, these sequences fortuitously allowed a comparison of the utility of these two regions for identifying *Dioryctria* species.

When using distance data to separate closely related species, it is particularly important to use the most informative region available. The variability in this data set suggests that the region of greatest divergence varies between taxa, which means that reliance on a single region within a gene could be misleading (e.g. the 475 bp region was most divergent in the *zimmermani* Gr., contrary to the general trend where the DNA barcoding region showed the greatest differences). This variability was also seen between species groups, particularly between the *abietella* Gr. and *auranticella* Gr.

Such variation in divergence could be due either to variation in the mutation rate or the level of constraint between these two fragments. Both of these processes could lead to differences in their observed substitution rate and their susceptibility to the effects of saturation due to multiple substitutions. Corrections, such as Kimura-2-parameter model (K2P) or LogDet can help mitigate the effects of multiple substitutions, although these corrections do not, in practice, fully eliminate the effects of saturation (Felsenstein 2004). The neighbor-joining method, favored by the DNA barcoding advocates (Hebert et al. 2003), is a distance method and may be significantly affected by saturation. A K2P correction of the pairwise distances in the current data set did not reduce the variability seen in pairwise comparisons (A.D.R. unpublished data). Substitution rate heterogeneity, secondary structure, mutation hot spots or even recombination could explain the divergence variability observed in my data set (Lunt et al. 1996, Hagelberg 2003, Howell et al. 2003, Doan et al. 2004, Ho et al. 2005). Variability between pairwise sequence divergence within these sequence fragments not only raises concerns regarding the use of a single region of mtDNA for predicting the presence of distinct species, it also contradicts any assumption of neutral or nearly neutral molecular evolution and provokes closer examination of the processes affecting mitochondrial DNA evolution.

Short fragments of COI (400 – 800 bp) have commonly been used to identify closely related species, particularly in Lepidoptera (Caterino et al. 2000). Although short fragments are used extensively and have been popularized for DNA barcoding (Hebert et al. 2003), such reliance on short COI fragments has been questioned (Wahlberg et al. 2003). Short fragments may have low numbers of phylogenetically informative characters, reducing their utility for separating closely related species. This effect is compounded when only a single specimen is used to define a lineage. For this reason, all major lineages recognized in this study contained at least three specimens, and, when possible, specimens from multiple populations (Table 2-1). These problems can be further reduced if longer DNA fragments are included in the analysis (Mitchell et al. 2000, Wahlberg et al. 2003, Wahlberg and Nylin 2003).

Although the 475 bp fragment resolved the eight species included in this study, mtDNA phylogenies represent a single genetic tree and therefore caution must be exercised when relying on mtDNA for delineating species. Cases where species trees and mtDNA gene trees are incongruent have been well documented, and they are especially common between the most closely related species (Avise and Ball 1990, Avise 1991, Nichols 2001, Funk and Omland 2003, Ballard and Whitlock 2004). To be confident that species delineated by the mtDNA data constitute unique biological entities (e.g., species), additional characters such as morphology, geographic range, and larval host plant should be examined.

Morphological Identification

Although morphological characters were generally sufficient to identify the lineages to species, the highly variable forewing pattern of some species was problematic. Specimens of *D. pentictonella* showed a wide range of forewing variation, ranging from pale, nearly white phenotypes to dark red and black phenotypes. Many of these phenotypes resembled other species in the region (e.g., *D. fordi* or *D. okanaganella*), particularly *D. ponderosae*. Although *D. ponderosae* was not collected during this study, it appears in species lists for California (Furniss and Carolin 1977, California Moth Specimen Database 2005). Caution must be exercised when relying on these identifications, due to the similarity between this species and some wing phenotypes of

D. pentictonella. Genitalic characters definitively separate these two species, as they are in separate species groups, but these characters are often not examined.

The phenotype RBR was similar to the original description for *D. pentictonella* (Mutuura et al. 1969b) and was found throughout the year, but several other wing phenotypes were also common (RBB and RWR; Table 2-3; Fig. 2-3). The wing phenotypes were compared to genetic variation, seasonality and pheromone lure attraction, but none of these factors accounted for the forewing pattern variation observed in the population (Tables 2-3 and 2-4; Fig. 2-3). Similar forewing variability was seen among the smaller number of *D. pentictonella* females. Forewing pattern in *D. pentictonella* thus appears to be highly variable and plastic, even within a single population. Such phenotypic plasticity is quite common in Lepidoptera, particularly with respect to forewing colour patterns. An extreme example of phenotypic forewing variation occurs in *Acleris cristana* (Denis & Schiffermüler) (Lepidoptera: Tortricidae) where 119 forms were described for a single species in Britain (Manley 1973). Understanding the morphological variability seen in *D. pentictonella* will require a detailed examination of the environmental and genetic factors influencing the development of wing pattern in this species.

Reared Material and Pheromone Attraction

MtDNA and morphological comparisons were essential for identifying *Dioryctria* species reared from host material or captured with pheromone traps. A single specimen of *D. auranticella* was captured in a pheromone trap (Table 2-4), and although considered a cone pest, no specimens were reared from ponderosa pine cones at the GRC even though it has previously been recorded on this host (Mutuura and Munroe 1972, Hedlin et al. 1980, Neunzig 2003). *Dioryctria auranticella* does not usually occur in large populations and often affects only a small percentage of cones in a given area (Hedlin et al. 1980), which may explain the lack of reared material. *Dioryctria cambiicola* specimens were reared from wounds on Douglas-fir and blister rust tissue on lodgepole pine. Although described from several species of pine, *D. cambiicola* has not been previously recorded on Douglas-fir.

The majority of specimens reared from cones were identified as *Dioryctria abietivorella* and, based on these numbers, likely caused the majority of cone damage at the GRC. *Dioryctria abietivorella* larvae are generalist feeders and have been reared from a wide variety of larval hosts (Table 2-1). Although large numbers of specimens were present in cones, relatively low numbers of specimens were captured in pheromone traps (Table 2-4), suggesting that the pheromone lures used in the study are only weakly attractive to *D. abietivorella*. Recent work on pheromone blends of *D. abietivorella* has found that (3Z, 6Z, 9Z, 12Z, 15Z) pentacosapentaene is a key component, and is attractive as a 5:1 ratio with (9Z,11E)-14:Ac (Millar et al. 2005) .

Dioryctria pentictonella, on the other hand, was captured in higher numbers by pheromone traps and three distinct flight periods were observed (Fig. 2-2). In contrast to the number of specimens captured in pheromone traps, only three specimens were reared from host material. *Dioryctria pentictonella* is primarily a foliage feeder, although it has occasionally been recorded on cones (Neunzig 2003). Collecting effort was targeted at species feeding within cones, rather than on foliage, which may explain the discrepancy between the numbers of adults captured by pheromone traps and those reared from host plants.

Multiple collecting methods were essential for sampling species in the study region, and mtDNA sequences allowed clear associations to be made among specimens collected by different methods, independently of morphological characters. Examination of reared material helped to identify the species causing the majority of cone damage at the GRC, but provided incomplete sampling of *Dioryctria* diversity in the area. Pheromone trapping provided insight into the attractiveness of different pheromone lures deployed to sample *Dioryctria* populations in the area, but was limited by the range of the blends deployed. Recent work has also shown that pheromone trapping may not accurately sample local populations, particularly in populations at the edges of ranges, and as a result may inaccurately evaluate genetic diversity and gene flow (Salvato et al. 2005). Most of the species, but not *D. pentictonella* at Chico, were sampled by UV or mercury vapor light trap (Table 2-1). Reliance on pheromone trapping, rearing or light trapping alone, therefore, would not have documented the diversity of species in the area. Collecting with a variety of methods across a broad temporal and geographic range

clearly provides the most accurate estimation of population structure and representation of species in the region.

A combination of mitochondrial, morphological and behavioural characters was needed to simplify identification of *Dioryctria* species in the study region. By sequencing mtDNA from specimens reared from cones, *D. abietivorella* was identified as the primary seed pest at the Genetic Resource Center in Chico, CA. This finding, in combination with clarifications of diagnostic morphological characters that identify this species, provides support for pest management. Likewise, the combination of these data demonstrated that *D. pentictonella* has a wide range of intraspecific forewing polymorphism. This species was trapped in large numbers by pheromone lures, but was not responsible for the majority of cone damage at the GRC. Using mtDNA to clarify morphologically difficult groups and improve pest identifications has proven very effective (Sperling et al. 1995, Kerdelhué et al. 2002, Scheffer et al. 2004, Simmons and Scheffer 2004). Identification of mtDNA lineages can allow detection and testing of morphological characters that more conveniently diagnose pest species and are correlated with identification on the basis of other characters. Contamination of molecular samples is always a concern, so having a suite of morphological characters to validate molecular identifications is essential. Morphology can more easily be examined in the field, and may not require laboratory facilities for identification of specimens. This speeds up identifications and improves the ability of forest managers to manage outbreaks, making morphological characters an important component of diagnostic keys and an essential complement to molecular studies.

Phylogenetic Relationships

Morphological examinations identified and assigned mtDNA lineages to eight *Dioryctria* species in six previously described species groups (Table 2-1). These species groups were originally described based on genitalic variation and forewing differences (Mutuura and Munroe 1972). Deeper phylogenetic relationships between lineages, particularly between species groups, were only poorly resolved by the 475 bp sequence data (Fig. 2-1). Consequently, I increased character sampling by sequencing the entire COI-COII region (2.3 kb) to resolve some of these relationships (Fig. 2-5).

Several interesting patterns were revealed when the phylogenetic relationships of the six *Dioryctria* species groups were examined in light of other characteristics (Fig. 2-5). For example, the *zimmermani* Gr. and *ponderosae* Gr., which were well supported as sister groups (Node F) in the molecular data, were also supported by several morphological and ecological synapomorphies. Larvae in this clade feed in the cambium of host plants (with the possible exception of *D. fordi* whose host is undetermined), often targetting wounds or blister rust infections and larval feeding causes the formation of pitch masses. Males have a constricted uncus and a valve with a hooked apical projection. Females have longitudinal wrinkles on the ductus bursa, though the size and depth of those wrinkles are variable.

The *baumhoferi* Gr., which had a moderately supported sister group relationship with the *zimmermani* Gr. + *ponderosae* Gr. clade (Node E; Fig. 2-5), was supported by two additional synapomorphies. All three species groups have raised scales in several regions of the forewing and form a monophyletic "raised-scale" group. The majority of species also specialize on pines, although exceptions do occur (e.g., *D. cambiicola*; Table 2-1). The species groups in this clade are the most speciose in North America north of Mexico, with 30 out of 40 described species classified under these three groups. Species of this clade each appear to specialize on only one or two *Pinus* species, rather than feeding on a wide range of host plants like *D. abietivorella* (Hedlin et al. 1980, Neunzig 2003). During the Tertiary, *Pinus* experienced a diversification throughout North America and is correlated to climatic change (Millar 1998). *Dioryctria* in the "raised-scale" clade may have radiated onto the pines during or sometime after their diversification leading to the present species diversity. No appropriate phycitine fossils are available to calibrate divergence rates for *Dioryctria*, and there is noticeable divergence rate heterogeneity among different species groups, so any hypothesis of evolution between *Dioryctria* and their larval hosts remains speculative.

The *schuetzeella* Gr. and *auranticella* Gr. formed a poorly supported clade (Node C). Specimens of this clade lacked raised scales, as did *D. abietivorella* and the outgroups, making this absence a plesiomorphic character. Species in this clade also lacked a constricted uncus and a prominent pre-apical spine, characters shared by members of the *baumhoferi* Gr. Larval host associations differ between the members of

these two clades, with the *schuetzeella* Gr. feeding in cones, foliage and cambium in a wide range of conifer species (Mutuura and Munroe 1973, Neunzig 2003), while species in the *auranticella* Gr. feed primarily in cones on *Pinus* species (Munroe 1959, Mutuura and Munroe 1972, Neunzig 2003).

The *abietella* Gr. was sister group to the rest of the *Dioryctria* species groups in the analysis of 2.3 kb sequences, but this relationship was poorly supported. *D. abietivorella* was the only representative of the species group and lacked raised scales like the previous two species groups. *D. abietivorella* feeds primarily on cones, like members of the *auranticella* Gr., although it will switch to foliage during times of low cone production (Trudel et al. 1999). This species is a generalist and feeds on a wide range of conifer hosts, unlike those in the "raised-scale" clade.

Increased character sampling, by examining the full COI-COII sequence, provided enough phylogenetically informative characters to develop a preliminary phylogeny for species and species groups of *Dioryctria*, but some parts of the phylogeny were still poorly supported. *Dioryctria* is a diverse genus and sampling of additional species is needed throughout North America, particularly from the speciose *zimmermani* Gr. and *baumhoferi* groups. Additional characters, such as from nuclear gene sequences or morphological analyses, will be needed to fully elucidate the relationships within this genus and to improve the resolution of the deeper clades in the phylogeny.

Table 2- 1: Locality and collection information for *Dioryctria* specimens used in mtDNA surveys in northern California, Oregon and British Columbia.

Locality Data	Collection[1]	Date[2]	No.	Collector, Year	Haplotype	GenBank Accession #	
						475 bp	2.3 kb
abietella Group							
D. abietivorella							
USA: CA: Butte Co., Chico	Pheromone – I, II, V	June – August	11	C. Rudolf, G. Grant (2000, 2001)	AB1	DQ296154	DQ295185
USA: CA: Butte Co., Chico	Pheromone trap	August – September	3	C. Rudolf, G. Grant (1998)	AB1		
USA: CA: Butte Co., Chico	Cone: Pp, Pb, Pl, Df		37	C. Rudolf (1995, 2000-2001)	AB1		
USA: CA: Butte Co., Chico	MV-light	June	8	A. Roe (2001)	AB1		
USA: CA: Butte Co., Chico	MV-light	June	1	A. Roe (2001)	AB4		DQ247740
USA: CA: Butte Co., Chico	MV-light	June	1	A. Roe (2001)	AB5		DQ247741
USA: CA: Placer Co., Foresthill	Cone: Pp		3		AB1		
USA: OR: Lane Co., Cottage Grove, Dorena Tree Center	Cone: Pm, Lo		3	J. Berdeen (2001)	AB1		
USA: OR: Clackamas Co., Colton, Horning Tree Center	Cone: Df		1	B. Willhite (2001)	AB1		
USA: CA: Butte Co., Chico	Cone: Pb, Pl, Df		6	C. Rudolf (2000-2001)	AB2	DQ296156	
USA: CA: Butte Co., Chico	MV-light	June	1	A. Roe (2001)	AB2		
USA: CA: Butte Co., Chico	Cone: Df		1	C. Rudolf (2001)	AB3	DQ296155	
auranticella Group							
D. auranticella							
USA: CA: El Dorado Co., Placerville	MV-light	June	2	A. Roe (2001)	OS1	DQ296157	DQ295176
USA: CA: El Dorado Co., Placerville	MV-light	June	1	A. Roe (2001)	OS4		DQ247736
USA: CA: Butte Co., Chico	MV-light	June	1	A. Roe (2001)	OS2		
USA: CA: Butte Co., Chico	Pheromone – II	June	1	C. Rudolf, G. Grant (2001)	OS2		
USA: CA: El Dorado Co., Placerville	MV-light	June	2	A. Roe (2001)	OS2	DQ296158	
D. rossi							
CAN: BC: 35 km E Summerland	MV-light	August	2	A. Roe (2003)	OS3	DQ296159	DQ295177

Table 2-1 cont

Locality Data	Collection1	Date2	No.	Collector, Year	Haplotype	GenBank Accession #	
						475 bp	2.3 kb
schuetzeella Group							
D. pseudotsugella							
USA: OR: Benton Co., Corvalis	UV-light	July	2	J. Adams (2001)	RE1	DQ296160	DQ295186
USA: OR: Benton Co., Corvalis	UV-light	July	1	J. Adams (2001)	RE2	DQ296161	
USA: OR: Benton Co., Corvalis	UV-light	July	1	J. Adams (2001)	RE3	DQ296162	
CAN: BC: 8 km E Adams Lake	MV-light	August	1	A. Roe (2003)	RE4	DQ296163	
CAN: BC: 10 km SW Pritchard	MV-light	August	1	A. Roe (2003)	RE5	DQ296164	
baumhoferi Group							
D. pentictonella							
USA: CA: Butte Co., Chico	Pheromone – V, III, IV	May – September	46	C. Rudolf, G. Grant (2000, 2001)	RS1a	DQ296165	DQ295180
USA: CA: Butte Co., Chico	Pheromone trap	April – September	6	C. Rudolf, G. Grant (1998)	RS1a		
USA: CA: Siskiyou Co., Ball Mt.	Cone: Pa		2	J. Stein (1994)	RS1a		
USA: CA: Butte Co., Chico	Cone: Pb		1	C. Rudolf (1997)	RS1a		
USA: CA: El Dorado Co. Placerville	MV-light	August	1	A. Roe (2001)	RS1a		
CAN: BC: 35 km E Summerland	MV-light	August	1	A. Roe (2003)	RS1a		
USA: CA: Butte Co., Chico	Pheromone – V, III	May - June	14	C. Rudolf, G. Grant (2000, 2001)	RS1b	DQ296166	
USA: CA: Butte Co., Chico	Pheromone trap	May, July	2	C. Rudolf, G. Grant (1998)	RS1b		
USA: CA: Alameda Co. Berkeley	UV-light	June	1	FAH Sperling (1998)	RS1b		
USA: CA: Butte Co., Chico	Pheromone trap	May	1	C. Rudolf, G. Grant (1998)	RS1c		DQ295181
USA: CA: El Dorado Co., Placerville	MV-light	June	1	A. Roe (2001)	RS1d		DQ295182
USA: CA: Butte Co., Chico	Pheromone - III	October	1	C. Rudolf, G. Grant (2000)	RS1e	DQ296167	
USA: CA: Siskiyou Co., Ball Mt.	UV-light	September	1	C. Frank	RS1f	DQ296168	
zimmermani Group							
D. cambiicola							
USA: OR: Medford	Bark: Df		1	J. Berdeen (2001)	RS2a	DQ296169	
CAN: BC: Prince George Tree Improvement Station	Bark: Pc		2	A. Roe (2001)	RS2b	DQ296170	DQ295183

Table 2-1 Cont.

Locality Data	Collection[1]	Date[2]	No.	Collector, Year	Haplotype	GenBank Accession #	
						475 bp	2.3 kb
D. fordi							
USA: CA: Butte Co., Chico	MV-light	June	1	A. Roe (2001)	RS2c		DQ295184
USA: CA: Butte Co., Chico	MV-light	October	2	A. Roe (2002)	RS2h	DQ296173	
USA: CA: Butte Co., Chico	MV-light	October	1	A. Roe (2002)	RS2g	DQ296174	
ponderosae Group							
D. okanaganella							
USA: CA: El Dorado Co., Placerville	MV-light	June	1	A. Roe (2001)	RS2d	DQ296171	
USA: CA: El Dorado Co., Placerville	MV-light	June	1	A. Roe (2001)	RS2e		DQ295179
USA: CA: El Dorado Co. Blodgett Research Stn. 15 mi E Georgetown	MV-light	August	1	A. Roe (2002)	RS2e	DQ296172	
USA: CA: El Dorado Co., Placerville	MV-light	June	1	A. Roe (2001)	RS2f		DQ295178
Outgroups							
Oncocera faecella							
China: Inner Mongolia: Mt. Manhan	Light	August	1	D. Zhang (2002)	O. fae		DQ247727
Ceroprepes ophthalmicella							
China: Henan Province: Mt. Baiyun	Light	July	1	X. Wang (2002)	C. oph.		DQ247728

[1]Host records from material reared to adult or extracted as larvae. Host abbreviations: Pb: Afghan pine (*Pinus brutia* ssp. *eldarica* (Medw.) Nahal); Pl: sugar pine (*Pinus lambertiana*); Pp: ponderosa pine (*Pinus ponderosa*); Pc: lodgepole pine (*Pinus contorta* Douglas ex Loudon); Pa: whitebark pine (*Pinus albicaulis* Engelmann); Pm: western white pine (*Pinus monticola* Douglas ex D. Don); Df: Douglas-fir (*Pseudotsuga menziesii*); Lo: western larch (*Larix occidentalis* Nuttall).
[2]Date is not given for reared material due to unreliability of emergence times in artificial conditions.

Table 2- 2: Mitochondrial DNA primers used in surveying *Dioryctria* spp. over 2.3 kb of COI-COII.

Primer Name	Direction and Location (3' end)[1]	Sequence (5' – 3')
K698	TY-J-1460	TAC AAT TTA TCG CCT AAA CTT CAG CC
RonV	C1-J-1751	GGA GCT CCA GAT ATA GCT TTC CC
K699	C1-N-1840	AGG AGG ATA AAC AGT TCA (C/T)CC
K808	C1-N-1840	TGG AGG GTA TAC TGT TCA ACC
Jerry*	C1-J-2183	CAA CAT TTA TTT TGA TTT TTT GG
JerryV*	C1-J-2183	CAA CAT TTA TTT TGA TTC TTT GG
Nancy	C1-N-2191	CCC GGT AAA ATT AAA ATA TAA ACT TC
K525	C1-N-2329	ACT GTA AAT ATA TGA TGA GCT CA
Brian	C1-J-2495	CTT CTA TAC TTT GAA GAT TAG G
MilaIII*	C1-N-2659	ACT AAT CCT GTG AAT AAA GG
George	C1-J-2792	ATA CCT CGA CGT TAT TCA GA
GeorgeIII	C1-J-2792	ATA CCT CGG CGA TAC TCT GA
GeorgeV	C1-J-2792	ATA CCT CGA CGA TAT TCC GA
PatII	TL2-N-3013	TCC ATT ACA TAT AAT CTG CCA TAT TAG
Pierre	C2-J-3138	AGA GCC TCT CCT TTA ATA GAA CA
Marilyn	C2-N-3389	TCA TAA GTT CA(A/G) TAT CAT TG
MarilynII	C2-N-3389	TCA TA(T/A) CTT CA(A/G) TAT CAT TG
MarilynIII	C2-N-3389	TCA TAT CTT CAG TAT CAC TG
Preston	C2-J-3570	GCA ACA GAT GTT ATT CAC TCT TG
Eva	C2-N-3782	GAG ACC ATT ACT TGC TTT CAG TCA TCT

[1]Following Simon et al. 1994: J/N: Majority/Minority (equivalent to sense/antisense for COI-COII).

*Primer combination used for the 475 bp fragment.

Table 2- 3: Wing phenotypes for *Dioryctria* specimens used in mtDNA survey.

Species	Haplotype	Wing Phenotype[1] (No.)
D. abietivorella	AB1	NBB (31), NBR (9), NBW (3), NWB (1), UNS (22)
	AB2	NBB (4), UNS (3)
	AB3	NBB (1)
	AB4	NBB (1)
	AB5	NBB (1)
D. auranticella	OS1	NOR (2)
	OS2	NOR (2), UNS (2)
	OS4	NOR (1)
D. rossi	OS3	NOR (2)
D. pseudotsugella	RE1	NBrR (2)
	RE2	UNS (1)
	RE3	NBrR (1)
	RE6	NBrR (1)
	RE7	NBrR (1)
D. pentictonella	RS1a	RWR (15), RBB (14), RBR (14), RWB (5), RWW (4), RBW (1), UNS (4)
	RS1b	RWR (6), RBR (4), RBB (1), RWB (3), RWW (2)
	RS1c	RBR (1)
	RS1d	RWR (1)
	RS1e	RWR (1)
	RS1f	UNS (1)
D. cambiicola	RS2a	UNS (1)
	RS2b	UNS (2)
D. fordi	RS2c	RWR (1)
	RS2g	RWR (1)
	RS2h	RWR (2)
D. okanaganella	RS2d	RBB (1)
	RS2e	RBB (1), UNS (1)
	RS2f	RBR (1)

[1] Code for wing phenotypes:
- 1st position: raised scales
  - N – absent
  - R – present
- 2nd position: primary forewing colour
  - B – black to dark grey
  - W – white to pale grey
  - O – orange
  - Br – brown to tan
- 3rd position: colour of subbasal area
  - B – black to dark grey
  - W – white to pale grey
  - R – tan to dark red

Table 2- 4: *Dioryctria* specimens collected at five pheromone blends, characterized for mtDNA haplotypes, and wing phenotypes (all from the Genetic Resource Center, Chico, CA, 2000 – 2001).

Species	Haplotype	No.	Blend	Components1 (Dose [µg])	Wing Phenotype (No.)
D. abietivorella	AB1	9	I	Z (100) + E (1)	NBR (4), NBB (2), NWB (1), UNS (2)
		1	II	T (100)	UNS
		1	V	T (100) + D (5)	NBR
D. auranticella	OS1	1	II	T (100)	NOR
D. pentictonella	RS1a	42	V	T (100) + D (5)	RWR (12), RBB (10), RBR (10), RWB (5), RWW (3), RBW (1), UNS (1)
		3	III	T (10) + D (0.5)	RBB, RBR, RWR
		1	IV	T (1) + D (0.05)	RWW
D. pentictonella	RS1b	12	V	T (100) + D (5)	RBR (4), RWB (3), RWR (2), RWW (2), RBB (1)
		2	III	T (10) + D (0.5)	RBB, RWR
D. pentictonella	RS1e	1	III	T (10) + D (0.5)	RWR

[1]Chemical names of *Dioryctria* pheromone blend components:

Z: (Z,E)-9,11-tetradecadienyl acetate
E: (Z,E)-9,12-tetradecadienyl acetate
T: (Z)-9-tetradecenyl acetate
D: (Z)-7-dodecenyl acetate

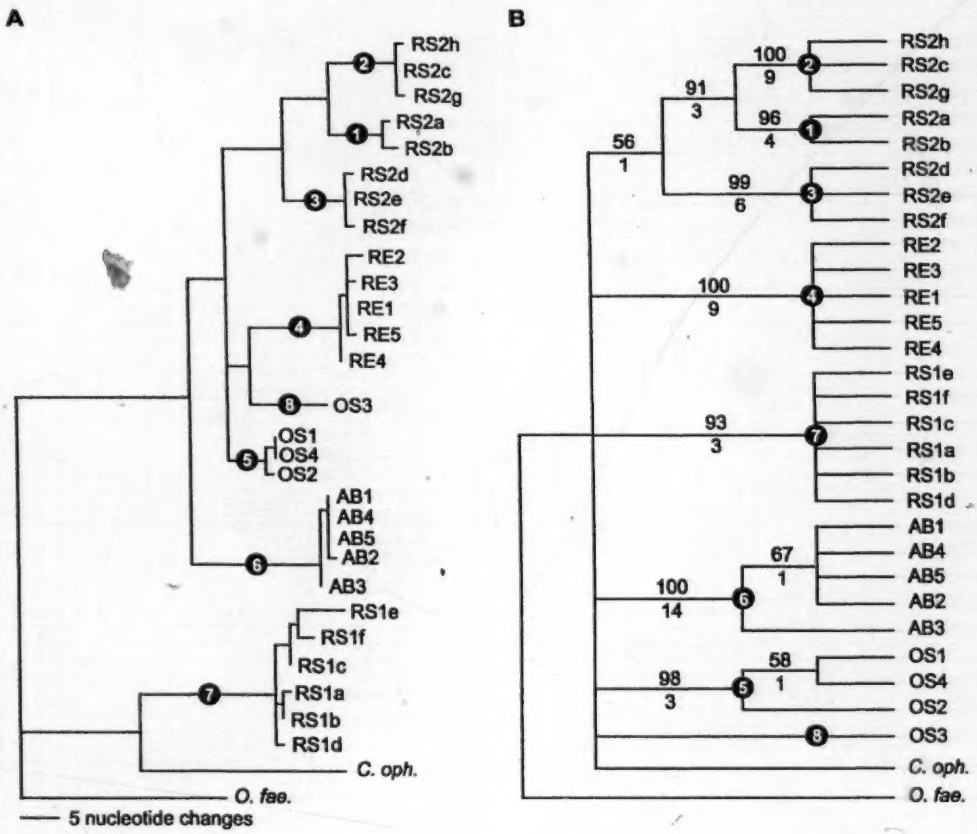

Figure 2- 1: Maximum parsimony analysis of 475 bp COI. (A) Phylogram of one of 178 most-parsimonious trees (length = 182 steps; CI = 0.692; RI = 0.862) showing phyletic branch lengths. Lineages labeled 1 – 8 are discussed in text. (B) Strict consensus of 178 most-parsimonious trees, with bootstrap values >50% shown above branches and Bremer support values below branches. Haplotype codes are explained in Table 2-1.

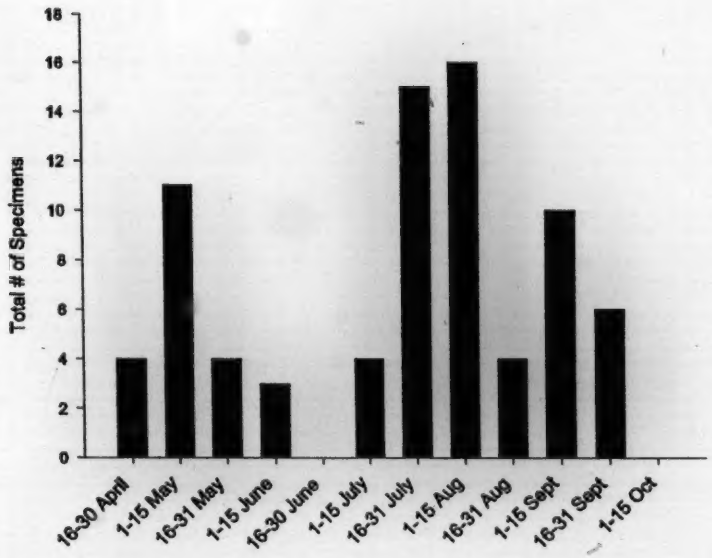

Figure 2- 2: Flight period for *D. pentictonella* based on all pheromone and light-trapped material included in study, including 1998 pheromone-trapped material.

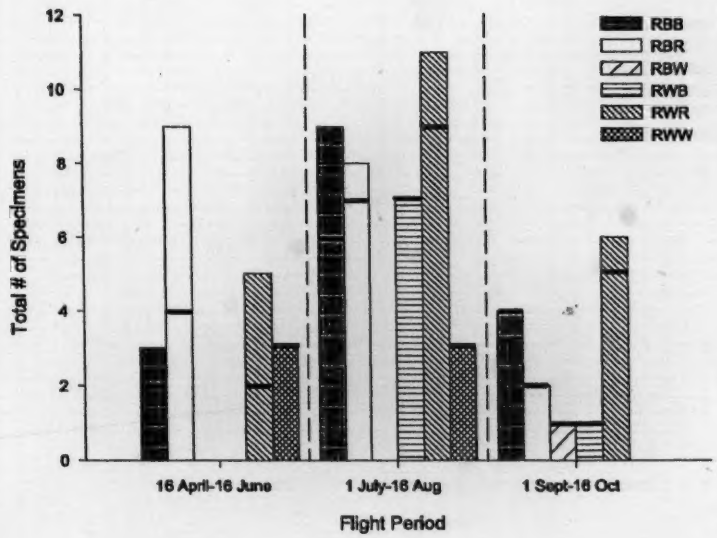

Figure 2- 3: Wing phenotypes collected during three flight periods of *D. pentictonella*. Dark horizontal bars separate number of specimens collected by pheromone traps (below) from other methods (above). Wing phenotype coding is discussed in text.

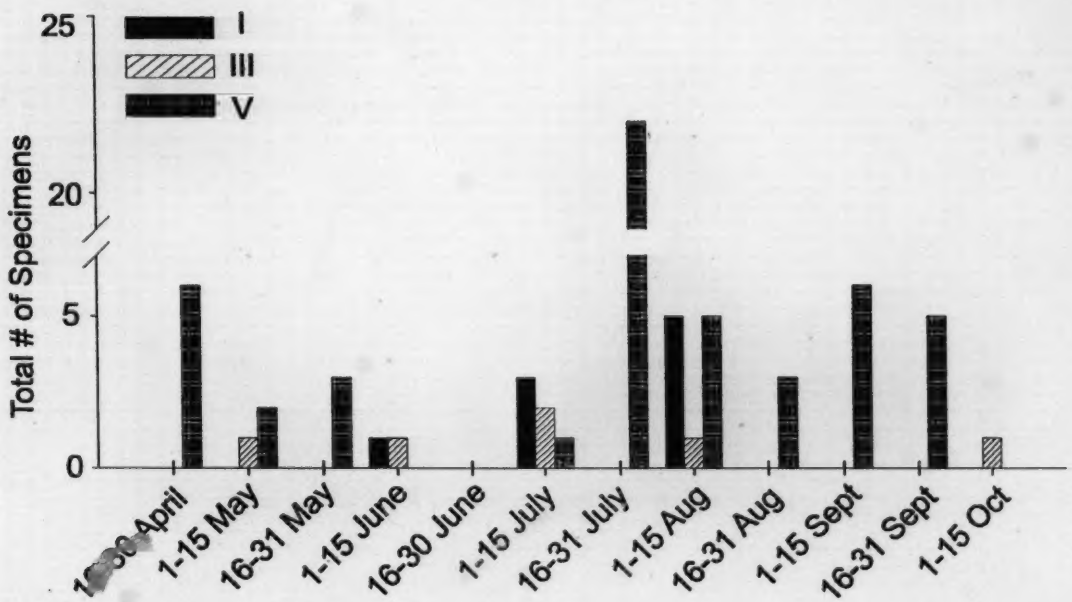

Figure 2- 4: Flight times for specimens collected from three pheromone lures (I, III, V; lure components explained in Table 2-4) from 2000 and 2001 which were also sequenced for mtDNA. Lures with catch totals of fewer than three specimens are not shown.

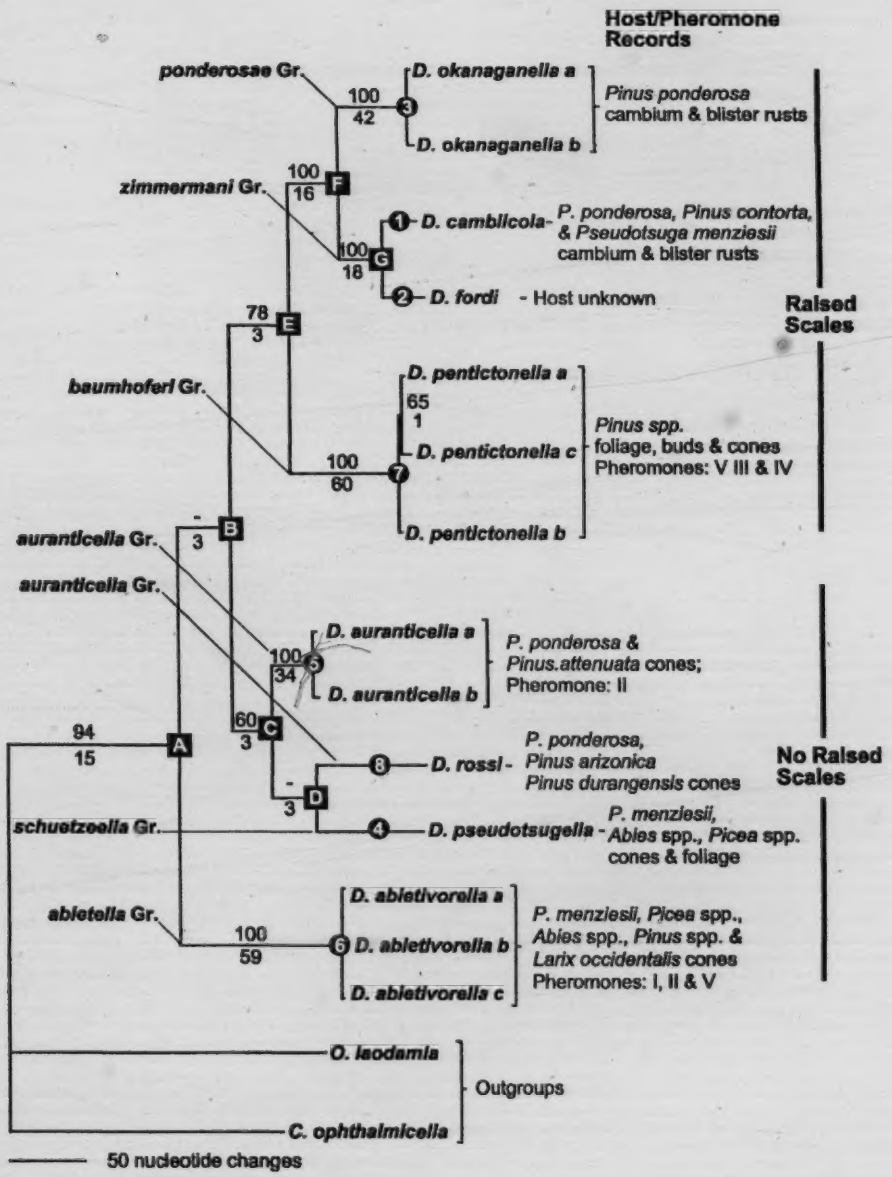

Figure 2- 5: Phylogram of single most parsimonious tree for 2.3 kb COI-COII (length = 840; CI = 0.714; RI = 0.746). Lineages labeled 1 – 8, and Nodes A – G are discussed in text. Known host associations and pheromone attraction are indicated for each species, Bootstrap values >50% are shown above branches and Bremer support values below branches.

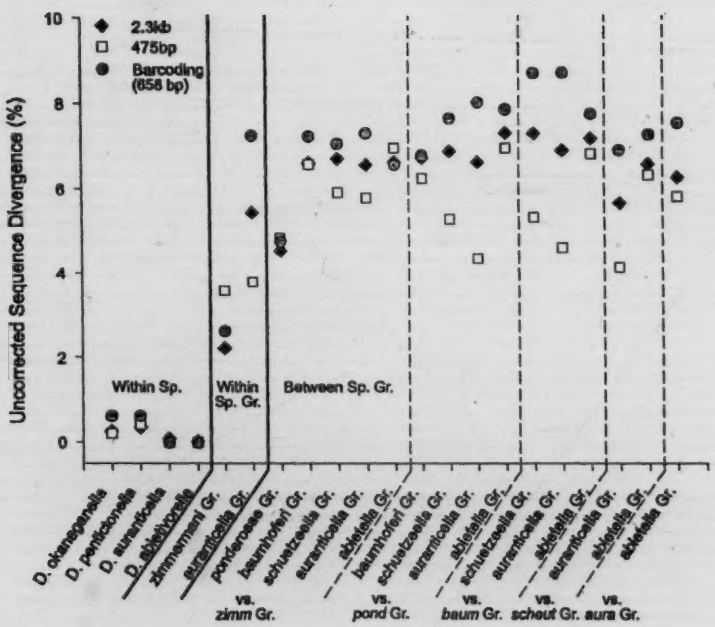

Figure 2- 6: Comparison of uncorrected sequence divergences within species, within species groups and between species groups for *Dioryctria* specimens sequenced across the full 2.3 kb of COI+COII. Abbreviations of between species groups comparisons are as follows: *zimm* Gr.= *zimmermani* Gr; *pond* Gr.= *ponderosae* Gr.

Literature Cited

Avise, J. C. 1991. Ten unorthodox perspectives on evolution prompted by comparative populations genetic findings on mitochondrial DNA. Annual Review of Genetics 25:45-69.

Avise, J. C., and Ball, R. M. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. Oxford Surveys in Evolutionary Biology 7: 45-67.

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13:729-744.

Blake, E. A., M. R. Wagner, and T. W. Koeber. 1989. Relative effect of seed and cone insects on ponderosa pine in northern Arizona. Journal of Economic Entomology. 82: 1691-1694.

Blanchard, A., and E. C. Knudson. 1983. A new species of *Dioryctria* Zeller (Lepidoptera: Pyralidae) from Texas. Proceedings of the Entomological Society of Washington 85: 116-120.

California Moth Specimen Database. 2005. http://elib.cs.berkeley.edu/eme/calmoth.html. Accessed June 28th, 2005.

Caterino, M. S., S. Cho, and F. A. H. Sperling. 2000. The current state of insect molecular systematics: A thriving tower of Babel. Annual Review of Entomology 45: 1-54.

Cibrián-Tovar, D., B. H. Ebel, H. O. Yates, and J. T. Méndez-Montiel. 1986. Insectos de conos y semillas de las coníferas de México/Cone and seed insects of the Mexican conifers. U.S. Dept. Agr. Gen. Tech. Rep. SE-40. 110 pp.

Clary, D. O., and D. R. Wolstenholme. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. Journal of Molecular Evolution 22: 252-271.

Coulson, R., A. Mutuura, and E. Munroe. 1972. The *Dioryctria* species of loblolly pine in east Texas, with comments on the occurrence of two new species. Journal of Economic Entomology 65: 868-870.

Donahue, J. P., and H. H. Neunzig. 2005. Two new *Dioryctria* from California (Lepidoptera: Pyralidae: Phycitinae). Holarctic Lepidoptera 10: 33-36.

Doan, J. W., T. R. Schmidt, D. E. Wildman, M. Uddin, A. Goldberg, M. Hüttemann, M. Goodman, M. L. Weiss, and L. I. Grossman. 2004. Coadaptive evolution in cytochrome *c* oxidase: 9 of 13 subunits show accelerated rates of nonsynonymous substitution in anthropoid primates. Molecular Phylogenetics and Evolution 33: 944-950.

Du, Y., A. D. Roe, and F. A. H. Sperling. Phylogenetic framework for *Dioryctria* Zeller (Lepidoptera: Pyralidae: Phycitinae) based on combined analysis of mitochondrial DNA and morphology. Canadian Entomologist (in press).

Eriksson T. 2002. AutoDecay, version 5.0.3. Distributed by author, Bergius Foundation, Royal Swedish Academy of Sciences, Stockholm, Sweden.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution and Systematics 34:397-423.

Felsenstein, J. 2004. Inferring Phylogenies. Sinauer Associates, Inc., Sunderland, MA.

Furniss, R. L., and V. M. Carolin. 1977. Western Forest Insects. U.S. Department of Agriculture Forest Service.

Grant, G. G., Y. H. Prevost, K. N. Slessor, G. G. King, and R. J. West. 1987. Identification of the sex pheromone of the spruce coneworm, *Dioryctria reniculelloides* (Lepidoptera: Pyralidae). Environmental Entomology 16: 905-909.

Grant, G. G., S. A. Katovich, D. J. Hall, D. A. Lombardo, and K. N. Slessor. 1993. Sex-pheromone identification and trapping of *Dioryctria resinosella* (Lepidoptera, Pyralidae). Environmental Entomology 22: 154-161.

Hagelberg, E. 2003. Recombination or mutation rate heterogeneity? Implications for mitochondrial eve. Trends in Ecology and Evolution 19: 84-90.

Hainze, J. H., and Benjamin, D. M. 1984. Impact of the red pine shoot moth, *Dioryctria resinosella* (Lepidoptera: Pyralidae), on height and radial growth in Wisconsin red pine plantations. Journal of Economic Entomology 77: 36-42.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London Series B 270: 313-322.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, W. Hallwachs. 2005. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator.* Proceedings of the National Academy of Science U. S. A. 101: 14812-14817.

Hedlin, A. F., H. O. Yates, D. Cibrián-Tovar, B. H. Ebel, T. W. Koerber, and E. P. Merkel. 1980. Cone and seed insects of North American conifers. 122 pp. Joint Publication: Environment Canada, Canadian Forest Service; Sec. Agr. Recursos Hidáulicos, Mexico.

Heinrich, C. 1956. American Moths of the Subfamily Phycitinae. Smithsonian Institution, Washington D.C.

Ho, S. Y. W., M. J. Phillips, A. Cooper, A. J. Drummond. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular Biology and Evolution 22. 1561-1568.

Howell, N., C. Bogolin Smejkal, D. A. Mackey, P. F. Chinnery, D. M. Turnbull, and C. Hernstadt. 2003. The pedigree rate of sequence divergence in the human mitochondrial genome: these is a difference between phylogenetic and pedigree rates. American Journal of Human Genetics 72: 659-670.

Kerdelhué, C., G. Roux-Morabito, J. Forichon, J. M. Chambon, A. Robert, and F. Lieutier. 2002. Population genetic structure of *Tomicus piniperda* L. (Curculionidae: Scolytinae) on different pine species and validation of *T. destruens* (Woll.). Molecular Ecology 11: 483-494.

Knölke, S., S. Erlacher, A. Hausmann, M. A. Miller, and A. Segerer. 2005. A procedure for combined genitalia dissection and DNA extraction in Lepidoptera. Insect Systematics and Evolution 35: 401-409.

Kruse, J. J., and F. A. H. Sperling. 2001. Molecular phylogeny within and between species of the *Archips argyrospila* complex (Lepidoptera: Tortricidae). Annals of the Entomological Society of America 94: 166-173.

Landry, B., J. A. Powell, and F. A. H. Sperling. 1999. Systematics of the *Argyotaenia franciscana* (Lepidoptera: Tortricidae) species group: evidence from mitochondrial DNA. Annals of the Entomological Society of America 92: 40-46.

Lipscomb D., N. Platnick, and Q. Wheeler. 2003. The intellectual content of taxonomy: a comment on DNA taxonomy. Trends in Ecology and Evolution 18: 65-66.

Löfstedt, C., J. N. C. Van der Pers, J. Lofqvist, and B. S. Lanne. 1983. Sex pheromone of the cone pyralid *Dioryctria abietella*. Entomologia Experimentalis et Applicata 34: 20-26.

Lunt, D., D.-X. Zhang, J. M. Szymura, and G. M. Hewitt. 1996. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. Insect Molecular Biology 5: 153-156.

Lyons, L. A. 1957. Insects affecting seed production in red pine. II. *Dioryctria disclusa* Heinrich, *D. abietella* (D. and S.) and *D. cambiicola* (Dyar) (Lepidoptera: Phycitidae). Canadian Entomologist 89: 70-79.

Manley, W. B. L. 1973. A guide to *Acleris cristana* [Denis & Schiffermüller, 1775] (Lepidoptera: Tortricidae) in Britain. Entomology Gazette 24: 89-206.

Meyer, W. L., G. L. DeBarr, C. W. Berisford, L. R. Barber, and W. L. Roelofs. 1982. Identification of the sex pheromone of the webbing coneworm moth, *Dioryctria disclusa* (Lepidoptera: Pyralidae). Environmental Entomology 11: 986-988.

Meyer, W. L., R. S. Cameron, A. Tamhankar, G. L. DeBarr, C. W. Berisford, and W. L. Roelofs. 1984. Sex pheromone of the blister coneworm moth, *Dioryctria clarioralis* (Lepidoptera: Pyralidae). Environmental Entomology 13: 854-858.

Meyer, W. L., G. L. DeBarr, J. L. Hanula, B. Kovalev, R. S. Cameron, C. W. Berisford, and W. L. Roelofs. 1986. (Z)-11-hexadecenyl acetate, a sex pheromone component for the southern pine coneworm, *Dioryctria amatella* (Lepidoptera: Pyralidae). Environmental Entomology 15: 316-320.

Millar, C. I. 1998. The early evolution of pines, pp. 69-91. *In*. D. M. Richardson (ed.) Ecology and biogeography of *Pinus*. Cambridge University Press. New York, NY.

Millar, J. G., G. G. Grant, J. Steven McElfresh, W. Strong, C. Rudolf, J. D. Stein, and J. A. Moreira. 2005. (3Z, 6Z, 9Z, 12Z, 15Z)-pentacosapentaene, a key pheromone component of the fir coneworm moth *Dioryctria abietivorella*. Journal of Chemical Ecology 31: 1229-1234.

Mitchell, A., C. Mitter, and J. C. Regier. 2000. More taxa or more characters revisisted: combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insects: Lepidoptera). Systematic Biology 49: 202-224.

Mosseler, A., B. A. Roberts, and P. Tricco. 1992. The effects of fir coneworm, *Dioryctria abietivorella* (Grote) (Lepidoptera: Pyralidae), on seed production in small isolated populations of red pine, *Pinus resinosa* Ait. Forest Ecology and Management 53: 15-27.

Munroe, E. 1959. Canadian species of *Dioryctria* Zeller (Lepidoptera: Pyralidae). Canadian Entomologist 91: 65-72.

Mutuura, A. 1982. American species of *Dioryctria* (Lepidoptera: Pyralidae) VI. A new species of *Dioryctria* from eastern Canada and north-eastern United States. Canadian Entomologist 114: 1069-1076.

Mutuura, A., and E. Munroe. 1972. American species of *Dioryctria* (Lepidoptera: Pyralidae) III. Grouping of species: Species of the *auranticella* group, including the Asian species, with the description of a new species. Canadian Entomologist 104: 609-625.

Mutuura, A., and E. Munroe. 1973. American species of *Dioryctria* (Lepidoptera: Pyralidae) IV. The *schuetzeella* group and the taxonomic status of the spruce cone moth. Canadian Entomologist 105: 653-668.

Mutuura, A., and E. Munroe. 1979. American species of *Dioryctria* (Lepidoptera: Pyralidae) V. Three new cone-feeding species from the southeastern United States. Journal of the Georgia Entomological Society 14: 290-304.

Mutuura, A., E. Munroe, and D. A. Ross. 1969a. American species of *Dioryctria* (Lepidoptera: Pyralidae) I. Western Canadian species of the *zimmermani* Group. Canadian Entomologist 101: 1009-1023.

Mutuura, A., E. Munroe, and D. A. Ross. 1969b. American species of *Dioryctria* (Lepidoptera: Pyralidae) II. Western Canadian species of the *baumhoferi* and *ponderosae* groups. Canadian Entomologist 101: 1042-1047.

Neunzig, H.H. 2003. Pyraloidea, Pyralidae (part), Phycitinae (part): R. B. Dominick, et al. The Moths of America North of Mexico, fasc. 15.5. The Wedge Entomological Research Foundation, National Museum of Natural History, Washington, DC.

Neunzig, H. H. 1990. A new species of *Dioryctria* (Pyralidae: Phycitinae) from Mexico. Proceedings of the Entomological Society of Washington 92: 493-496.

Neunzig, H. H., and N. A. Leidy. 1989. A new species of *Dioryctria* (Lepidoptera: Pyralidae: Phycitinae) from the southeastern United States. Proceedings of the Entomological Society of Washington 91: 321-324.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology and Evolution 16: 358-364.

Paquin, P. and M. Hedin. 2004. The power and perils of 'molecular taxonomy': a case study of the eyeless and endangered *Cicurina* (Araneae: Dictynidae) from Texas caves. Molecular Ecology 13: 3239-3255.

Richmond, J.A. 1995. Macrogeographic genetic variation in populations for the webbing coneworm (Lepidoptera: Pyralidea). Journal of Entomological Science 30: 349-358.

Richmond, J.A. and M. Page. 1995. Genetic and biochemical similarities among four species of pine coneworms (Lepidoptera: Pyralidae). Annals of the Entomological Society of America 88: 271-280.

Salvato, P., M. Simonato, L. Zane, T. Patarnello, L. Masutti, and A. Battisti. 2005. Do sexual pheromone traps provide biased information of the local gene pool in the pine processionary moth? Agricultural and Forest Entomology 7: 127-132.

Schaber, B. D., and F. E. Wood. 1971. A new species of *Dioryctria* infesting loblolly pine. Proceedings of the Entomological Society of Washington 73: 215 - 223.

Scheffer, S. J., R. M. Giblin-Davis, G. S. Taylor, K. A. Davies, M. Purcell, M. L. Lewis, J. Goolsby, and T. D. Center. 2004. Phylogenetic relationships, species limits, and host specificity of gall-forming *Fergusonina* flies (Diptera:

Fergusoninidae) feeding on *Melaleuca* (Myrtaceae). Annals of the Entomological Society of America 97: 1216-1221.

Schowalter, T. D., M. I. Haverty, and T. W. Koeber. 1985. Cone and seed insects in Douglas-fir, *Pseudotsuga menziesii* (MIRB.) Franco, seed orchards in the western United States: Distribution and relative impact. Canadian Entomologist 117: 1223-1230.

Simmons, R. B., and S. J. Scheffer. 2004. Evidence of cryptic species within the pest *Copitarsia decolora* (Guenée) (Lepidoptera: Noctuidae). Annals of the Entomological Society of America 97: 675-680.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87: 651-701.

Sopow, S. L., R. G. Bennet, J.-F. Landry, and B. Landry. 1996. Identification of the "grey' *Dioryctria* species of British Columbia (Lepidoptera: Pyralidae). Journal of the Entomological Society of British Columbia 93: 75-91.

Sperling, F. 2003a. Butterfly Molecular Systematics: From species definitions to higher-level phylogenies, pp. 431-458. *In* C. Boggs, W. Watt and P. Ehrlrich [eds.], Butterflies: Evolution and Ecology Taking Flight. University of Chicago, Chicago Press, Chicago.

Sperling, F. 2003b. DNA barcoding: Deus ex machina. Newsletter of the Biological Survey of Canada (Terrestrial Arthropods) 22: 50-53.

Sperling. F. A. H., and D. A. Hickey. 1994. Mitochondrial DNA sequence variation in the spruce budworm species complex (*Choristoneura*: Lepidoptera). Molecular Biology and Evolution 11: 656-665.

Sperling, F. A. H., J.-F. Landry, and D. A. Hickey. 1995. DNA-based identification of introduced ermine moth species in North America (Lepidoptera: Yponomeutidae). Annals of the Entomological Society of America 88: 155-162.

Sperling, F. A. H., A. G. Raske, and I. S. Otvos. 1999. Mitochondrial DNA sequence variation among populations and host races of *Lambdina fiscellaria* (Gn.) (Lepidoptera: Geometridae). Insect Molecular Biology 8: 97-106.

Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony. Version 4 beta 10. Sinauer Associates, Sunderland, MA.

Trudel, R., E. Bauce, C. Guertin, and J. Cabana. 1999. Performance of the fir coneworm *Dioryctria abietivorella* (Grote) as affected by host species and presence or absence of seed cones. Agricultural and Forest Entomology 1: 189-194.

Wahlberg, N., and S. Nylin. 2003. Morphology versus molecules: resolution of the positions of *Nymphalis*, *Polygonia,* and related genera (Lepidoptera: Nymphalidae). Cladistics. 19: 213-223.

Wahlberg, N., R. Oliveira, and J. A. Scott. 2003. Phylogenetic relationships of *Phyciodes* butterfly species (Lepidoptera: Nymphalidae): complex mtDNA variation and species delimitations. Systematic Entomology 28: 257-273.

Will, K., and D. Rubinoff. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics. 20: 47-55.

Winter, W. D. 2000. Basic Techniques for Observing and Studying Moths and Butterflies. Memoirs of the Lepidoterists' Society No. 5. 444 pgs.

Chapter 3: Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for species recognition

Introduction

Molecular systematics is one of the most rapidly expanding fields in biology and yet our understanding of patterns of molecular character evolution remains relatively superficial. Analysis of molecular data has proven to be important for understanding deep phylogenetic relationships (Blair and Hedges 2005, Edwards et al. 2005, Regier et al. 2005), identifying morphologically similar immatures (Olson 1991, Wells et al. 2001, Sharley et al. 2004), examining population structure within a species (Avise et al. 1987, Zhang and Hewitt 2003), assigning unknown specimens to reference species (Bartlett and Davidson 1992, Baker and Palumbi 1994, Hebert et al. 2003a, b), and diagnosing and delimiting cryptic species (Goetze 2003, Scheffer et al. 2004, Blair et al. 2005, Hebert et al. 2005, Hendrixson and Bond 2005). However, delimiting species using only molecular characters has always been controversial (Dunn 2003, Lipscomb et al. 2003, Seberg et al. 2003, Tautz et al. 2003), and recent debates have been particularly heated (Moritz and Cicero 2004, Ebach and Holdrege 2005, Hebert and Gregory 2005, Meyer and Paulay 2005, Will et al. 2005) in light of the increasing popularity of projects such as the Consortium for the Barcode of Life (http://barcoding.si.edu/).

Studies that use DNA fragments to delimit species, either separately or in combination with morphological and ecological data, have relied on a variety of gene markers. Gene choice and fragment length vary substantially, depending on the evolutionary question and taxon of interest (Caterino et al. 2000, Meyer and Zardoya 2003). MtDNA genes have long dominated the field of molecular systematics because of their maternal inheritance, limited recombination, rapid evolution, and the robustness of mtDNA against degradation, making them ideal markers for many species-level questions (Avise et al. 1987). Species-level vertebrate studies have focused primarily on cytochrome b or the control region (Sheldon and Bledsoe 1993, Honeycutt et al. 1995), while arthropod research has focused on a range of genes, in particular COI, COII, ND5

or 16S (Caterino et al. 2000). Fragment location within these gene regions is also variable, often dependent on the taxonomic group as well as the availability of previously developed primers (Folmer et al. 1994, Simon et al. 1994).

One of the key features of the DNA barcoding project, as proposed by Hebert et al. (2003a), is the designation of a single mtDNA fragment at the 5' end of cytochrome *c* oxidase I (COI) gene to act as a 'barcode' to identify and delineate all animal life. By choosing a standard DNA fragment, the efforts of multiple research groups can be coordinated, and they are able to construct a more compatible library of DNA sequences than would be possible if working independently (Caterino et al. 2000).

Although many aspects of DNA barcoding have been critiqued (Sperling 2003, Moritz and Cicero 2004, Will and Rubinoff 2004, Ebach and Holdrege 2005, Will et al. 2005), there has been only limited discussion on the decision to use a specific ~600 bp fragment from the 5' end of COI as the DNA barcoding region (Erpenbeck et al. 2005). Furthermore, there has been little discussion of the potential effects of using this particular region on delimiting closely related species. Initially, the 5' end of COI was chosen as the focal region because it is flanked by two "universal" primers that work for a range of metazoans (Folmer et al. 1994). The need to use robust primers is understandable, but examination of the DNA barcoding literature reveals that the majority of projects actually rely on taxon specific primers, rather than universal primers, in order to optimize PCR performance (Hebert et al. 2004, Penton et al. 2004, Barrett and Hebert 2005, Hebert et al. 2005), particularly with degraded material (Lambert et al. 2005). In addition, some DNA barcoding projects have used even smaller fragments of COI (Whiteman et al. 2004, Page et al. 2005). With current sequencing technology, more than 900 bp are routinely obtained with 98.5% accuracy from a single run (Gunning et al. 2002), so it is not necessary to limit the DNA fragment length to 600 bp. As well, maximum amounts of evidence should ideally be examined when inferring species boundaries (Dayrat 2005, Will et al. 2005), due to stochastic variation in the genome. Thus it is reasonable to expect projects that rely solely on shorter DNA fragments for delimiting species to be more vulnerable to heterogeneous patterns of nucleotide substitutions within COI. The key question is whether there is a region or length that is optimally informative.

Unequal substitution patterns, heterogeneous evolutionary rates, and functional constraints can potentially all affect the usefulness of COI for species delimitation. The COI gene is composed of interspersed, highly conserved membrane bound regions and variable extramembrane loops (Fig. 3-1) (Lunt et al. 1996). The COII gene is composed of several distinct sections: N-terminal region, membrane bound regions, copper-binding region, linking region and carboxy terminus (Fig. 3-1) (Saraste 1990). Previous examinations of COI molecular evolution have shown that the structure and function of this gene can impact the accumulation of mutations (Saraste 1990, Lunt et al. 1996, Caterino and Sperling 1999), although the majority of studies have focused on amino acid substitution rate, rather than nucleotide change. Nucleotide differences between closely related species generally occur at the third codon position, and patterns of nucleotide substitution are likely to be overlooked if only amino acid changes are compared. Mutation hot spots, or adaptive substitutions, are known to exist in mtDNA (Stoneking 2000, Innan and Nordborg 2002). These processes may play an important role in shaping phylogenetic relationships and population structure, and may vary between independent lineages and change as taxa become increasingly diverged (Galtier et al. 2006). I expected that regions of high divergence, and consequently saturation, will be good indicators of phylogenetic signal, and may play an important role in our ability to detect phylogenetic relationships and population structure. I have focused on both the COI and COII genes because many projects have relied on both (Caterino et al. 2000) and the existence of sequences that cover the full 2.3kb COI+COII region allowed me to make more tightly controlled comparisons.

Considering the growth of DNA barcoding and DNA taxonomy (Monaghan et al. 2006), it is urgent to understand how patterns of nucleotide substitution within COI-COII can affect the delimitation of closely related species. It is also important to examine whether previous work has used the most informative gene regions. A variety of criteria could be considered when choosing an optimally informative COI-COII region. Such a region should maximize one or more of nucleotide divergence, robustness of phylogenetic signal, consistency of evolutionary rate, or accuracy as an indicator of total genomic differences. Ideally, a region should be chosen to maximize several of these factors simultaneously, although some tradeoffs will be inevitable.

In this study I compare intra- and interspecific divergences among multiple groups of closely related COI-COII sequences to examine variation in 1) localized nucleotide divergence patterns, 2) relative accumulated divergence, and 3) phylogenetic signal. In particular, I examine how the particular COI fragment used by the DNA barcoding project compares to other similarly sized fragments throughout these genes. Finally, I provide recommendations for a more accurate and effective DNA region for delimiting closely related species.

Methods

Overview

In order to visualize patterns of nucleotide divergence throughout COI-COII, I chose to use sliding window analysis (Rozas et al. 2003). I calculated a specific parameter (e.g. nucleotide divergence) using a preset window and step size (e.g. 600 bp window with 5 bp step), with each window represented by the median nucleotide location, or nucleotide midpoint. From this, a graphical representation or profile of nucleotide divergence along a length of sequences is produced. I compared nucleotide diversity within species and nucleotide divergence between sister species pairs across full COI-COII sequences. Profiles often contained regions of concentrated nucleotide divergence, suggesting the possibility of mutational hot spots. Using randomly generated sequences, I examined whether empirically observed regions of high divergence exceeded the stochastic variation produced by simulated random divergence, which would provide evidence for biologically significant hot spots. I expressed nucleotide divergence in each window relative to total COI-COII divergence, allowing me to combine and compare all sister species pair profiles. I also examined how fragment length and location affected relative divergence, and identified COI-COII sequence regions that could act as accurate indicators of total COI-COII nucleotide divergence. Finally, sliding window analysis was used to examine the relationship between nucleotide divergence and the frequency of transitions, which were used as a measure of phylogenetic signal along the length of COI-COII.

Available Material

A total of 114 mtDNA sequences were used in this analysis (Appendix 1). These represent all available insect sequences (as of December 2005) that would allow comparison of haplotypes within and between closely related species. These sequences contiguously spanned a 2.3 kb region of mtDNA that begins in the tRNA tyrosine gene, crosses the COI, tRNA leucine, and COII genes, and ends in the tRNA lysine gene. The DNA barcoding region of Hebert et al. (2003a) includes the 5' half of COI. All sequences were aligned by eye to the sequence of *Drosophila yakuba* (accession number: NC 001322) and nucleotide position was numbered based on the *D. yakuba* system (*Dy*#1460 to *Dy*#3774 Clary and Wolstenholme 1985). The aligned data set contained 2342 bp, although sequence length varied between species, ranging from 2254-2320 bp. Length variation was restricted to the 5' end of COI, and within the tRNA-Leu region.

A total of 73 taxa (species and sub-species; See Appendix 1) were examined from Lepidoptera and Diptera; no addtional groups of complete COI-COII insect sequences were available outside these two orders. Eighteen sequences were previously unpublished (Appendix 1) and four previously published partial sequences were completed for the full COI-COII region. These were obtained following DNA extraction, PCR amplification and sequencing protocols outlined in Sperling and Hickey (1994). Heterologous primers designed for a range of insect families were used to develop primers for each new insect species (Bogdanowicz et al. 1993, Simon et al. 1994, Sperling and Hickey 1994). New sequences were deposited in GenBank under accession numbers DQ792576-DQ792593, while extended sequences retained the original accession numbers, and the remaining 92 sequences were already available on GenBank. All accession numbers and associated publications are listed in Appendix 1.

To understand the relationship between nucleotide divergence patterns and species delimitation, patterns of change within species, as well as between recognized sister species must be examined (Moritz and Cicero 2004). It is at this taxonomic level that identifications and delineations are most likely to be problematic. A total of 23 species with multiple COI-COII sequences (Table 3-1) were available to examine intraspecific variability. I also examined 23 pairs of sister species for interspecific

variability (Table 3-2). Sister species pairs were chosen based on previously established phylogenetic relationships (see associated literature in Appendix 1).

Sliding Window Analysis

Patterns of nucleotide substitution across COI-COII were visualized using sliding window analyses performed with DNAsp ver. 4.10.4 (Rozas et al. 2003). Nucleotide diversity (π) (Nei 1987 equation 10.5, Rozas et al. 2003) was used to calculate variability within species. Nucleotide divergence (K) (Tajima 1983 equation A3, Rozas et al. 2003) was used to analogously quantify variability between sister species. Both values were converted to percentages to facilitate comparisons. Analyses were performed using default settings (except for relative divergence profiles, see below), which included a Jukes-Cantor correction to all divergence calculations. To ensure that these results could be compared to data obtained from DNA barcoding, as well as other commonly used regions in COI-COII, I selected a 600 bp window size for all sliding window analyses. The universal primers used by Hebert et al. (2003a) (Folmer et al. 1994) produce a 658 bp fragment (*Dy*#1515-2172), while the original sequences deposited in GenBank from the initial DNA barcoding paper were either 617 or 624 bp in length (*Dy*#1556-2172 bp to *Dy*#1556-2179 bp, respectively) and subsequent studies have used variable lengths of sequence (Whiteman et al. 2004, Page et al. 2005). Three additional universal primer regions were also compared: C1-J-2183 to TL2-N-3014 (Jerry-Pat), C1-J-2792 to C2-N-3389 (George-Marilyn), and C2-J-3183 to TK-N-3775 (Pierre-Eva) (Fig. 3-1) (Bogdanowicz et al. 1993, Simon et al. 1994). Actual region length (Table 3-3) is variable, so a 600 bp fragment from the middle of each region was used for comparison (similar to the barcoding region). Ranges and midpoints (in brackets) of each fragment are as follows: Jerry-Pat, *Dy*#2299-2899 (2599); George-Marilyn *Dy*#2790-3390 (3090); Pierre-Eva *Dy*#3134-3734 (3434).

A step size of 5 bp was used for all sliding window analyses, with each window represented by its nucleotide midpoint. A 5 bp step size was chosen to provide a smoothing function to the sliding window analysis, which improved visualization of nucleotide changes along the COI-COII fragment. Using the default setting, sites with indels were ignored and every window included 600 nucleotides. This ensured that

windows were comparable across profiles of COI-COII. To discern patterns of nucleotide change, I compared diversity profiles within species and divergence profiles between sister species pairs in order to identify 1) magnitude and locations of maximum divergence, 2) magnitude and locations of minimum divergence, and 3) magnitude of divergence in the four described primer regions (Fig. 3-1).

Randomizations

I identified regions of high nucleotide divergence in some profiles of sister species pairs, that could be considered mutational hot spots (Stoneking 2000, Galtier et al. 2006). I wished to evaluate whether these regions of high divergence were non-random peaks of divergence. To test for the presence of non-random peaks of divergence, I compared a sliding window profile to a distribution of randomly generated sequences simulated from the original profile. If the original profile significantly exceeds the distribution of the generated profiles, the peak of divergence could be considered non-random. The *Papilio* sequences comprised the most complete data set available in this study and the parameters from these were used to generate the random data sets. A likelihood tree (log-likelihood score –12699.587) was obtained from a maximum likelihood analysis of the original *Papilio* sequences using the following GTR+Γ+I model: base frequencies A=0.3255, C=0.1014, G=0.1206, T=0.4525; rates AC=10.6213, AG=16.7683, AT=8.8273, CG=1.5416, CT=122.9118, GT=1.000; gamma shape parameter=0.8468, and proportion of invariable sites =0.5750. Modeltest (Posada and Crandall 1998) was initially used to identify the most appropriate model for the data set using a hierarchical likelihood ratios test. A total of 50 random replicates were then generated in Mesquite (Maddison and Maddison 2004) based on the likelihood tree using the above GTR model and a starting branch depth of 0.1. A sliding window analysis was performed on three simulated sequence pairs that matched the nucleotide divergence of three real sister species pairs (*Papilio canadensis* x *P. glaucus*; *P. demodocus* x *P. erithonioides*; *P. troilus* x *P. palamedes*). Nucleotide divergence profiles in the random sequences were compared to the original data sets to identify any regions of divergence that fell beyond the 95% confidence limits of the mean (n=50) random divergence.

Localized Nucleotide Divergence as an Indicator of Total COI-COII Divergence

To assess whether a smaller region of sequence within COI-COII could act as an effective indicator of total COI-COII divergence, I first needed to convert absolute nucleotide divergence in each window into a form that would allow comparisons between species pairs with different total nucleotide divergences. This was accomplished by calculating the ratio of nucleotide divergence sampled in a window relative to total COI-COII divergence for that sister species pair, allowing me to combine all species pairs. Mean divergence was then calculated for each window for all 23 sister species pairs and combined into a single profile using the following option in DNAsp: Sliding Window Options: Sites with alignment gaps considered. Mean relative divergence (presented as a percentage) and standard deviation were calculated for each window.

I used two criteria to identify regions that would act as the best indicators for total COI-COII divergence. First, mean relative divergence for the region should be 100% of total COI-COII divergence. Second, regions should have a low standard deviation from mean relative divergence, indicating that there is minimal variation among sister species pairs. I also compared the relative total COI-COII divergence for the four commonly used primer regions described previously. I calculated mean absolute nucleotide divergence and divergence relative to COI-COII totals for each region for all 23 sister species pairs and identified regions that were effective indicators of total COI-COII divergence based on both mean and minimal variance.

Fragment Length

In addition, I examined how increasing fragment length affected estimates of total COI-COII divergence in the 23 sister species comparisons. Relative divergence was calculated as described above for each fragment length for each sister species pair. Starting with a 200 bp fragment at the 5' end of COI genes, fragment length was increased by 200 bp increments until the full COI-COII sequence data was included. This process was repeated starting at the 3' end of COII, as well as in the middle of COI-COII (midpoint *Dy*#2610, fragment length increased 100 bp in either direction). For each set, I calculated mean divergence relative to total COI-COII nucleotide divergence and

standard deviation for each fragment length and identified a length where these two parameters leveled off relative to total COI-COII.

Phylogenetic Signal

In addition to using absolute and relative divergences to identify hot spots, I also examined how phylogenetic signal changed across COI-COII. Molecular characters that are initially used for identifications or diagnoses are often later incorporated into studies with a broader phylogenetic scope. Consequently, it is important to know how well phylogenetic signal is retained across COI-COII at increasingly deeper divergences.

Phylogenetic signal gradually accumulates in the form of mutations as lineages diverge and, in a simplistic sense, greater divergences indicate a greater amount of phylogenetic information. However, at the same time, the probability of multiple hits at a given site increases, causing phylogenetic signal to degrade in the process called saturation. Transition/transversion ratio (Ti/Tv ratio) can be used as an indirect measure of saturation since this ratio changes predictably with saturation. Transitions occur much more commonly than transversions at low levels of divergence, particularly in mtDNA, but as two sequences diverge from a common ancestor, transversions gradually increase, obscuring previous changes (Galtier et al. 2006). As the Ti/Tv ratio approaches 1:2 in sequences with equal base frequencies, phylogenetic signal is lost, reducing the informativeness of a DNA sequence comparison. Phylogenetic signal can thus be considered a tradeoff between overall divergence and Ti/Tv ratio, with different relative amounts of these two quantities necessitating different kinds of phylogenetic analyses (Swofford et al. 1996).

I measured percent transitions (%Ti), rather than Ti/Tv ratio, to eliminate the problem of undefined values when transversions equal zero. Three series of species were compared for changes in %Ti, in individual phylogenetic time series. One reference taxon in each series was compared to a set of increasingly diverged congeners to examine how %Ti and absolute nucleotide divergence changed with increasing phylogenetic depth. For *Papilio machaon oregonius* the increasingly more distant comparisons were with *P. m. hippocrates*, *P. hospiton*, *P. indra*, and *P. xuthus*. For *P. canadensis* the comparisons were with *P. glaucus*, *P. rutulus*, *P. multicaudatus*, and *P. scamander*, all of

which are placed in a different subgenus than the series that includes *P. machaon oregonius*. For *Dioryctria zimmermani*, which is a pyralid moth, the comparisons were with *D. fordi*, *D. okanaganella*, and *D. magnifica*. A sliding window analysis of nucleotide divergence was conducted for each phylogenetic time series (600 bp window; 5 bp step), with Ti/Tv ratio being calculated for each window using PAUP* (vers 4.10b). Ti/Tv ratio was then converted to %Ti and graphed together with percent nucleotide divergence for each pair within the three phylogenetic time series. Regions of low %Ti were defined as regions where %Ti was less than or equal to 50% (Ti/Tv $\leq$ 1).

Results

Intraspecific Diversity

Sliding window profiles differed substantially among the 23 species that had multiple COI-COII sequences, and 12 profiles are presented here to illustrate the range of this diversity (Fig 3-2). Several species had more than one region of maximum diversity (Table 3-1: *Dioryctia pentictonella, Anopheles gambiae, D. sylvestrella,* and *Choristoneura fumiferana*). Although the location of maximum diversity was variable, some sliding window profiles bore substantial similarity to one another (Fig. 3-2: *Papilio demoleus* and *Papilio machaon*).

Full COI-COII diversity within species ranged from 2.71% (*P. demoleus*) to 0.029% (*D. abietivorella, H. robertsi peplidis*) with a mean diversity of 0.68% (Table 3-1). Maximum diversity among all 600 bp windows within the profiles ranged from 4.50% (*P. demoleus*) to 0.11% (*Dioryctria abietivorella, Hyles robertsi peplidis*) with a mean of 1.12% (Table 3-1). Minimum diversity among the windows ranged from 1.67% (*Feltia jaculifera*) to 0.0% (12 species) with a mean of 0.26% (Table 3-1).

Mean diversity within species for the four fragments was highest in the barcode fragment (0.84%) and lowest for the George-Marilyn fragment (0.46%) (Table 3-1). DNA barcoding region diversity ranged from 2.16% (*F. jaculifera*) to 0.0% (*C. fumiferana, D. abietivorella, Dioryctria auranticella, D. sylvestrella, H. robertsi peplidis*) with a mean of 0.84% (Table 3-1). Jerry-Pat region diversity ranged from 3.22% (*Feltia jaculifera*) to 0.00% (*Dioryctria abietivorella, D. auranticella, Hyles robertsi peplidis,*

and *Chrysomya megacephala*) with a mean diversity of 0.76% (±0.96) (Table 3-1). George-Marilyn region diversity ranged from 1.83% (*Papilio demoleus*) to 0.00% (*Dioryctria abietivorella, D. sylvestrella, Hyles robertsi peplidis,* and *Chrysomya megacephala*) with a mean diversity of 0.46% (±0.54) (Table 3-1). Pierre-Eva region diversity ranged from 2.11% (*Feltia jaculifera*) to 0.00% (*Dioryctria pentictonella* and *Chrysomya megacephala*) with a mean diversity of 0.58% (±0.59) (Table 3-1).

Many differences consisted of single nucleotide changes and the locations of peak diversity occurred throughout COI-COII (Fig. 3-2), but five species had regions of maximum divergence that contained midpoint *Dy#* 2042 (fragment *Dy#* 1743-2342) (Table 3-1). By comparison, only two species had a region of maximum divergence that corresponded to the DNA barcoding region (midpoint *Dy#* 1856, fragment *Dy#* 1556-2156). As well, at low levels of diversity (one or two nucleotide differences), regions of maximum diversity were often found at the 3' end of COII, while at higher levels of diversity the regions of maximum diversity were more frequently found in COI (Table 3-1)

Interspecific Divergence

Profiles of divergence between the 23 sister species pairs examined in this study were quite variable; 15 were chosen to illustrate the range of this variation (Fig. 3-3). Several species pairs had multiple distinct (≥ 100 bp apart) regions of maximum divergence (Table 3-2: *Papilio thoas* x *P. cresphontes, Papilio zelicaon* x *P. polyxenes, Dioryctria reniculelloides* x *D. pseudotsugella*). As was the case for diversity within species, divergence profiles were highly variable, but some sister pair profiles were remarkably similar (*Papilio demodocus* x *P. erithonioides* and *Papilio troilus* x *P. palamedes*), even between distantly related lineages (*Chrysomya megacephala* x *C. bezziana,* and *Hyles annei* x *H. euphorbiarum*) (Fig. 3-3).

Interspecific divergences for the full COI-COII region ranged widely, from 6.67% (*Papilio troilus* x *P. palamedes*) to 0.18% (*Yponemeuta cagnagella* x *Y. padella*), with a mean nucleotide divergence of 2.74% (Table 3-2). Maximum divergences ranged from 9.02% (*Papilio troilus* x *P. palamedes*) to 0.50% (*Yponemeuta cagnagella* x *Y. padella*), with a mean of 3.93% (Table 3-2). Minimum divergences ranged from 5.00% (*Papilio*

troilus x *P. palamedes*, *Papilio thoas* x *P. cresphontes*) to 0.00% (*Yponemeuta cagnagella* x *Y. padella*) with a mean of 1.71% (Table 3-2). Divergence in the DNA barcoding region ranged from 6.98% (*Chrysomya norrisi* x *C. variceps*) to 0.33% (*Anopheles gambiae* x *A. arabiensis*) with a mean divergence of 2.79% (Table 3-2). Nucleotide divergence even varied by up to 4.02% within a single profile (Table 3-2: *Papilio troilus* x *P. palamedes*).

Although regions of maximum divergence occurred throughout COI-COII, four sister species pairs had regions of maximum divergence that contained nucleotide midpoint *Dy*# 2662 (fragment *Dy*# 2362-2962), and 8 additional pairs were within 100 bp of this location (Table 3-2). By comparison, only one sister pair (*Yponomeuta cagnagella* x *Y. padella)* had a region of maximum divergence that corresponded to the DNA barcoding region (midpoint *Dy*# 1856, fragment *Dy*# 1556-2156), and four additional pairs were within 100 bp of this midpoint (*Anopheles melas* x *A. quadriannulatus, Dioryctria reniculelloides* x *D. pseudotsugella*, *Papilio canadensis* x *P. glaucus*, and *Papilio zelicaon* x *P. polyxenes*) (Table 3-2). A single sister pair had multiple regions of maximum divergence that encompassed both midpoints (*Papilio zelicaon* x *P. polyxenes*).

When nucleotide divergence in the DNA barcoding region was compared with the region of maximum divergence, I found a surprising amount of difference between the two regions (Table 3-2). Differences ranged from 3.66% (*Chrysomya megacephala* x *C. bezziana*) to 0.00% (*Yponomeuta cagnagella* x *Y. padella)*. Nearly half (10 sister pairs) showed more than 1% difference and 5 pairs differed by more than 2% (*Chrysomya megacephala* x *C. bezziana*, *Papilio demodocus* x *P. erithonioides, Papilio troilus* x *P. palamedes, Papilio thoas* x *P. cresphontes, Choristoneura fumiferama* x *C.* α+β). For two sister pairs (*Anopheles gambiae* x *A. arabiensis* and *Dioryctria zimmermani* x *Dioryctria cambiicola*) the difference between the two regions was greater than their total COI-COII divergence.

Overlap of Intra- and Interspecific Divergence

Where possible, I compared values of intraspecific diversity and interspecific divergence, and for the majority of sister pairs these values did not overlap (Tables 3-1

and 3-2, e.g.: maximum intraspecific diversity of *Hyles galli* compared to minimum interspecific divergence between *H. gallii* x *H. nicaea*). One notable exception was between the intraspecific diversity of *Anopheles gambiae* and the interspecific divergence of *A. gambiae* x *A. arabiensis* (Fig. 3-4). When intraspecific diversity of *A. gambiae* was plotted with the interspecific divergence between *A. gambiae* x *A. arabiensis*, there were several locations were intraspecific diversity in *A. gambiae* exceed the interspecific divergence between *A. gambiae* and *A. arabiensis* (Fig. 3-4).

Randomizations

When the profiles of the *Papilio* sister species pairs (*Papilio canadensis* x *P. glaucus; Papilio demodocus* x *P. erithonioides; Papilio troilus* x *P. palamedes*) were compared to randomly generated sister species pair profiles, the real nucleotide divergence was within the distribution produced by the random data sets (Fig. 3-5: *Papilio troilus* x *P. palamedes*). Two of the three distributions (*Papilio canadensis* x *P. glaucus,* and *Papilio demodocus* x *P. erithonoide*; data not shown) were entirely within the 95% confidence interval of the random profiles. Only *Papilio troilus* x *P. palamedes* had a section that was outside the 95% confidence limits of the mean nucleotide divergence of the simulated sequences; it was nonetheless within the full set of random profiles (Figs. 3-5), suggesting that accumulation of nucleotide divergence in the examined profiles may be solely due to random nucleotide variability.

Chimeric Sequences

During the course of this investigation, I discovered that a sliding window analysis can identify chimeric sequences. I identified an erroneous chimeric sequence in *Papilio demodocus* (AY457588, Zakharov et al. 2004). The COII gene region was highly divergent and identical to a species of *Papilio demoleus malaynus* from Malaysia (accession number AF044000). I examined the original data set and discovered that an error had occurred when the sequence was submitted to GenBank. The corrected sequence file has been deposited on GenBank.

Localized divergence relative to total COI-COII divergence

Relative divergence was quite variable among windows across the entire sequence length (Fig. 3-6). Mean relative divergence for each window ranged from 124.5% (midpoints 2715 and 2720 bp) to 86.7% (3483 bp), with standard deviation ranging from ± 54.5% (1869 bp) to ± 14.6% (2133 bp). If a mean relative divergence equal to 100% and a low standard deviation indicate an optimal region, then the best 600 bp window spanned *Dy#* 1831-2430 (midpoint *Dy#* 2133 bp), with a mean percent divergence of 100.7% and standard deviation ± 14.6% (Fig. 3-6).

I also related patterns of nucleotide divergence to the values for the true length of commonly used PCR regions rather than the central 600 bp of these regions (Tables 3-1 and 3-2; Fig. 3-6). Mean absolute divergence for these fragments ranged from 2.40% (George-Marilyn) to 3.24% (Jerry-Pat), and variance ranged from ± 1.76 (LCO-HCO) to 2.22% (Jerry-Pat) (Table 3-3). Difference of mean relative divergence from total COI-COII divergence (100%) ranged from 11.8% (LCO-HCO) to 14.4% (Jerry-Pat), while variance ranged from ±26.42% (George-Marilyn) to ± 44.45% (LCO-HCO) (Table 3-3). Jerry-Pat had the highest absolute divergence, while LCO-HCO had the best relative divergence, and LCO-HCO had the best absolute variance, while George-Marilyn had the best relative variance. Overall, no single PCR fragment appears to optimally sample the sister pair divergences of this study; none outperform the optimal 600 bp fragment identified previously (*Dy#* 1831-2430), even though LCO-HCO and Jerry-Pat each span parts of this region.

Fragment Length

By increasing fragment lengths in 200 bp increments, I found that the expected approach to 100% and 0% was quite gradual for relative divergence and standard deviation (Fig 3-7). Decrease of mean relative divergence was not uniform for the three starting locations, and highly variable. Starting at the 5'end of COI, the initial mean relative divergence was 77%, but rapidly approached 100% once fragment length increased to 400 bp. Interestingly, mean relative divergence continued to increase as fragment length increased, indicating that a region of high relative divergence was being sampled. Mean relative divergence did not consistently remain near 100% until fragment

length was 2000 bp. Starting at the 3' end of COII, the initial mean relative divergence was 75%, similar to the 5' end (Fig. 3-7). Mean divergence decreased again when the fragment reached 800 bp, and did not approach 100% until fragment length was 1400 bp. As fragment length increased mean divergence did not deviate much from 100%, unlike the 5' start. Starting in the middle of COI-COII (*Dy*# 2610), the initial mean relative divergence for 200 bp was high (120%), much higher than the two other starting locations (Fig. 3-7). Mean divergence increased further at 400 bp, then decreased gradually to 100% at 1400 bp. Like the 5' start, mean divergence deviated again from a 100% mean at longer fragment lengths. Changes in variance were also not uniform among the three start locations (Fig. 3-7). Difference in variance and rate of change was most pronounced at small fragment lengths, but once fragment length reached 1400 bp decreases in variance were similar for all start locations. At a fragment length of 1000 bp, the variance for the midpoint start was less than either other start locations, and remained lower than the other two start locations for all larger fragment lengths.

Phylogenetic Signal

When nucleotide divergence and %Ti profiles for three phylogenetic time series were graphed together, several patterns emerged. First, %Ti was initially high between closely related species and decreased as nucleotide divergence increased. Second, regions of high divergence did not always correspond to regions of low %Ti. For example, *Papilio machaon oregonius* x *P. m. hippocrates* (Fig. 3-8: arrow at midpoint 2060 bp) or *Papilio canadensis* x *P. multicaudatus* (Fig. 3-8: midpoint 3510 bp) show areas where high nucleotide divergence and relatively high %Ti coincide. Optimal regions of maximum divergence and %Ti were variable across the genes and taxa, and there appeared to be no common optimal region in the series of species sampled. Third, accumulation of %Ti varied between series, and even within series, depending on the species comparisons. In both *P. machaon oregonius* and *P. canadensis* time series (Fig. 3-8), regions of low %Ti (less than 50%) occurred initially between closely related species (*Papilio machaon oregonius* x *P. m. hippocrates* and *Papilio canadensis* x *P. glaucus*), but were not persistent, and were recovered again when more diverged taxa were examined (*Papilio machaon oregonius* x *P. indra* and *Papilio canadensis* x *P.*

multicaudatus). On the other hand, in the *Dioryctria zimmermani* time series regions of low %Ti do not occur, even at relatively high levels of divergence (Fig. 3-8: *Dioryctria zimmermani* x *D. magnifica*). In the *Papilio canadensis* series Fig. 3-8, regions of low %Ti occur initially (1.28%), but then do not occur again, even at 7.39% divergence.

Discussion

Nucleotide divergence has been a primary criterion for delimiting species and detecting cryptic species in many initiatives such as DNA taxonomy (Tautz et al. 2003, Monaghan et al. 2006) and DNA barcoding (Hebert et al. 2003a). I have shown that nucleotide change is heterogenous throughout COI-COII. This finding contrasts with the report by Hebert et al. (2003b) that "sequence divergences in the halves [of COI] were closely similar" and "because of this congruence, the measures of sequence divergence for other species pairs are analysed without reference to their source region in the gene." Such arbitrary reliance on a single mtDNA fragment without compensation for the underlying heterogeneity of evolution of the molecular sequences could result in under- or overestimating species diversity. Four specific nucleotide divergence patterns were examined that could affect species delimitation and diagnosis: 1) mutation hot spots, 2) overlap of intra- and interspecific divergences, 3) levels of divergence relative to fragment size and location, and 4) saturated phylogenetic signal.

Mutation Hot Spots

Sliding window analyses demonstrated that nucleotide diversity and divergence were quite variable across COI-COII (Figs. 3-2 and 3-3), both within profiles and in comparisions among taxa. However, regions of high divergence (e.g. Fig. 3-3: *Papilio troilus* x *P. palamedes*) did not usually exceed the 95% confidence interval of the mean divergence between randomly generated sequence pairs (Fig. 3-5). Although the presence of mutational hot spots was not confirmed, I found several patterns that suggested that COI-COII is not undergoing purely random nucleotide change. First, both mean relative divergence and variance in some regions of COI-COII were consistently higher than others, rather than being equal throughout COI-COII (Fig. 3-6). Second,

patterns of nucleotide divergence showed similarity between species pairs, even between distantly related taxa (Fig. 3-3: *Chrysomya megacephala* x *C. bezziana,* and *Hyles annei* x *H. euphobarium*). Third, regions of concentrated nucleotide change were similar when overall divergence levels were equal (e.g. Table 3-1: also many species with very low intraspecific diversity had nucleotide changes restricted to the 3' end of COII). Finally, maximum divergence values for a number of taxa appear concentrated in a region encompassing 2362-2962bp (Table 3-2).

Non-random variation has previously been shown to occur in mtDNA, due to selection, nucleotide or codon bias, or functional constraints (Lin and Danforth 2003, Bartolome et al. 2005, Galtier et al. 2006). To further explore this phenomenon, conspecific specimens should be examined throughout their geographic ranges, to maximize sampling of mtDNA haplotype diversity, and across as large a sequence region as possible, to minimize the effect of localized mutational anomalies. Variability of divergence rates is particularly important in groups characterized by extremely high levels of intraspecific variation (Johnson et al. 2002) or low levels of interspecific divergence (Sperling et al. 1995).

Intra- versus Interspecific Divergence

When relying on molecular data to delimit species, it is important that nucleotide diversity within species does not overlap with divergence between species. There was major overlap between intraspecific diversity (Table 3-1: 0.029-2.71%) and interspecific divergence (Table 3-2: 0.18-6.76%) among all COI-COII sequences. The DNA barcoding region showed a similar range of overlap (Table 3-1: intraspecific: 0.00-3.33%; Table 3-2: interspecific: 0.33-6.98%). The majority of species with both intra- and interspecific divergence showed little to no overlap between these two values, with the exception of *Anopheles gambiae* x *A. arabiensis* (Fig. 3-4). Historically, these species have been notoriously difficult to separate, forming species flocks and hybridizing freely (Besansky et al. 2003). This example demonstrates that overlap of intra- and interspecific variation can pose significant problems for DNA barcoding projects. I do note that the DNA barcoding region had no overlap of intra- and interspecific variation for these two species (Tables 3-1 and 3-2). Nonetheless, this

example illustrates a potential problem in relying on a single short gene region to delimit closely related species.

These data also provide an opportunity to evaluate the accuracy of standard divergence thresholds for delimiting species. Such thresholds rely on the presence of genetic "breaks" (Hebert et al. 2003b) or "barcoding gaps" (Meyer and Paulay 2005) between intra- and interspecific divergences. In response to Hebert et al. (2004), Moritz and Cicero (2004) pointed out the importance of testing thresholds against recognized sister species, rather than a random sample of arbitrarily chosen relatives. As shown in *A. gambiae x A. arabiensis,* intra- and interspecific overlap can vary by location *within* COI-COII for a single sister pair, but how accurate are thresholds for the entire data set? All comparisons within this study were based on closely related sister taxa, where identifications are most likely to be problematic. If I apply a 3% threshold (Hebert et al. 2003a) to divergences obtained from the barcoding region in Lepidoptera, only 9 out of 19 sister pairs (47%) are correctly delimited (Table 3-2). If a more conservative threshold of 2% is used for the entire data set, 17 out of 23 sister pairs (71%) are correctly delimited (Table 3-2). Even the most conservative thresholds failed to achieve the results claimed by barcoding advocates (98% correct delimitation, Hebert et al. 2003a). Ranges of intraspecific diversity and interspecific divergence were quite broad (Tables 3-1 and 3-2), even considering the possibility that some very high levels of intraspecific diversity may indicate undescribed species and some very low levels of interspecific divergence could result from poor taxonomic classification.

Fragment Size and Location

Fragment choice can influence species diagnoses that rely solely on nucleotide divergence, particularly when based on short DNA fragments. If divergences are generally low between sister species, targeting regions of maximum divergence would ensure the greatest probability of consistently delimiting these taxa by obtaining the regions with informative nucleotide variation. However, to achieve maximum compatibility with other studies, it is advantageous to focus on regions that give accurate and consistent estimates of divergences across larger mtDNA regions. I identified several regions that have a mean nucleotide divergence for sister species pairs similar to

total COI-COII divergence, although these regions differ substantially in variability (Fig. 3-6). I identified a 600 bp fragment (*Dy*# 1831-2430) with a mean percent divergence of 100.7% relative to total COI-COII divergence, which also had relatively low variance (± 14.6% st. dev.). Based on my data, this 600 bp fragment would be the best indicator of total COI-COII divergence for sister species.

In recognition of the value of incorporating previously published data (Caterino et al. 2000), I focused on regions of COI-COII that are commonly used in insect systematics (Caterino et al. 2000, Hebert et al. 2003a), with their most widely used primer pairs for reference (Bogdanowicz et al. 1993, Folmer et al. 1994, Simon et al. 1994). If an ideal region for species diagnoses is considered to minimize variance and maximize the approach of mean relative nucleotide divergence to 100%, then regions may be compared on the basis of these criteria. None of the commonly used PCR fragments optimized these values to the extent shown by the 600 bp window demarcated by *Dy*# 1831-2430, although LCO-HCO and Jerry-Pat partially span this region. I recommend that future taxonomic projects consider lengthening commonly used primer regions to incorporate this optimal region.

I have demonstrated that locations of maximum divergence occur throughout COI-COII, and that the PCR fragments traditionally used for insect systematics had high levels of variation in nucleotide divergence (Table 3-3). I also demonstrated that peaks of divergence are similar to divergences obtained from randomly generated sequences (Fig. 3-5). Consequently I feel that it is more important to maximize fragment length, than to target specific regions within COI-COII. For example if primer region LCO-HCO (DNA barcoding region) were extended 300 bp in the 3' direction, it would contain the optimal 600 bp fragment (*Dy*# 1831-2430) identified in this study. This 900 bp fragment could still be sequenced as a single fragment according to Gunning et al. (2002).

Using a larger frament would also help to minimize the nucleotide variability caused by random variation in COI-COII (Fig 3-5). Based on Fig. 3-7, variance within each fragment region does not level off in all three sets until the fragment reaches at least 1400 bp in length. To ensure that a fragment length will produce accurate results across a range of taxa, minimizing the variance within a fragment is important. Small fragments

are more likely to be skewed by localized regions of unusual nucleotide divergence, whereas increased fragment length would reduce this risk.

Phylogenetic Signal

Several systematically important patterns of phylogenetic signal were observed in the three phylogenetic time series. Although %Ti generally decreased as nucleotide divergence accumulated, areas of maximum divergence often did not correspond to the regions of low %Ti (Fig. 3-8: *Papilio machaon oregonius* x *P. hospiton* or *Papilio canadensis* x *P. multicaudatus*). This trend contradicts my initial expectations that there should be a tradeoff between these two values. For closely related species, it is plausible to assume that saturation is not significant, but it is interesting to note that some areas of low divergence nonetheless already appear to show some saturation (Fig. 3-8: *Papilio canadensis* x *P. multicaudatus*). I also found similar signal patterns of low levels of divergence that appear to correspond to tRNA Leu which has different functional constraints, allowing the accumulation of transversions at low levels of divergence (Fig. 3-8: *Papilio machaon oregonius* x *P. m. hippocrates* and *Papilio canadensis* x *P. glaucus*).

Hebert and Gregory (2005) have stated that projects such as DNA barcoding are not reconstructing phylogenetic relationships but instead focus explicitly on species delimitation and diagnostics. On the other hand, it was stated that COI "is more likely to provide deeper phylogenetic insights than alternative [genes]" (pg 1931 Hebert et al. 2003b), and as such, many researchers who employ a DNA barcoding approach will use the data to examine relationships among their focal species. For phylogenetic informativeness, a standard region should maximize nucleotide divergence while minimizing saturation (measured here as %Ti). Nucleotide divergence and %Ti accumulated unequally along COI-COII sequences; among the three time series no optimal region for maximizing nucleotide divergence and %Ti was evident. This variability suggests that researchers wishing to optimize the use of molecular data for inferring phylogenetic relationships need to examine the underlying patterns of phylogenetic signal accumulation for their taxon of interest.

Conclusion

Over the past decade, molecular sequence data, in conjunction with morphological and ecological characters, have become integral components of the 'tool box' (Armstrong and Ball 2005) for systematics and taxonomy. As molecular methods become more accessible, increasing numbers of researchers are incorporating, or relying solely on, molecular characters to help clarify species problems or identify specimens. The use of molecular characters will only increase, particularly as the number of DNA barcoding projects continues to grow. These types of projects can have significant impacts on systematics, taxonomy, conservation, and pest identification. The danger of relying solely on molecular characters for delimiting and identifying species is that the results may be misleading, especially when relying upon a single DNA fragment. Consequently I offer several recommendations for future projects using COI-COII fragments to delimit and diagnose species.

First, I suggest that researchers should maximize the length of the DNA sequence used for initial pilot studies on any taxon. This will increase the chance of sampling labile areas of elevated nucleotide divergence and phylogenetic signal. Longer fragments will also minimize variation across taxa and be more likely to reflect broader patterns of nucleotide divergence. The importance of larger fragments and more characters has now been highlighted elsewhere (Erpenbeck et al. 2005), although many projects, such as DNA barcoding, are still opting to use only short DNA fragments. As a standard for barcoding, the Database Working Group of the Consortium of the Barcode of Life (http://barcoding.si.edu/DNABarCoding.htm) has proposed that all barcodes (unless otherwise approved) be a 648 bp region from the COI gene, starting at position 58 of the mouse mitochondrial genome, and containing at least 500 unambiguous bases (Hanner 2005). Assigning a standard sequence region was a particularly important step for the DNA barcoding initiative and follows the recommendations of Caterino et al. (2000). Since a large compilation of sequences is now available, it is unreasonable to recommend using an entirely independent COI fragment for the standard sequence, even though a potentially better region exists. Rather I suggest that sequence length requirements

should be expanded. With developments of new and more efficient sequencing it is routinely possible to obtain longer fragments.

Second, in-depth projects that rely on mtDNA divergences to delimit species (e.g. Monaghan et al. 2006) could benefit from targeting regions of maximum divergence, particularly for species characterized by low levels of divergence. I did not find a single optimal region of maximal divergence across all taxa; this region varied considerably even between closely related taxa. If fragment length is increased in pilot studies, then consistency can be preserved for identification across broader taxonomic ranges, and these can be followed by more extensive surveys using shorter fragments that maximize the survey efficiency for that taxon.

Third, in keeping with the recommendations of Funk and Omland (2003), I suggest that multiple specimens from across the known geographic range of species should be examined to obtain a more accurate estimation of intraspecific and interspecific nucleotide variability. With numerous sequences for each comparison, overlap of these values can be accurately identified.

Fourth, I recommend that researchers wishing to use DNA sequence data for both species recognition and phylogenetic inference be aware of the underlying variability in phylogenetic signal, and adjust their choice of DNA region to minimize saturation and maximize divergence for their taxon of interest.

As the field of systematics evolves, and reliance on molecular data becomes increasingly prevalent, understanding the patterns of evolution in molecular characters becomes increasingly important. Researchers must be aware of the patterns of character change occurring in their data set and consciously consider the effects of these patterns on species delimitation, diagnoses and phylogenetic inference. There are no simple, universal solutions to the full range of problems that systematists routinely deal with, and so my final recommendation is that molecular systematists strive to retain flexibility and nuance in their responses to these challenges.

Table 3- 1: Intraspecific nucleotide diversity (%) for 23 species with multiple specimens (N), and are calculated as the average diversity between all specimens within a species. Full COI-COII diversity (Total), maximum diversity (Max, 600 bp fragment), and minimum diversity (Min, 600 bp fragment) are shown. Midpoints are shown in *D. yakuba* numbers, and regions with multiple equal divergence values were combined when less than 100 bp apart. Diversity was also examined for representative 600 bp fragments for the DNA barcoding region (BC, midpoint *Dy*# 1856), Jerry-Pat (J-P, midpoint *Dy*# 2599), George-Marilyn (G-M, midpoint *Dy*# 3090), and Pierre-Eva (P-E, midpoint *Dy*# 3434).

Species[c]	N	Total%	Max%	Midpoint(s)	Min%	Midpoint(s)	BC	J-P	G-M	P-E
Lepidoptera										
Choristoneura α+β lineages	9	1.59	2.07	2470-2480	0.98	3187-3212	1.78	1.76	1.24	1.45
Choristoneura fumiferana	2	0.09	0.17	2150-2799, 2940-3534	0.00	1765-2145, 2750-2935	0.00	0.17	0.17	0.17
Dioryctria abietivorella	3	0.03	0.11	3134-3473	0.00	1759-3138	0.00	0.00	0.00	0.11
Dioryctria auranticella	2	0.09	0.33	3343-3408	0.00	1761-3138	0.00	0.00	0.17	0.17
Dioryctria okanaganella	2	0.26	0.67	1760-1880[a]	0.00	2710-2890	0.67	0.17	0.17	0.17
Dioryctria pentictonella	3	0.35	0.67	1766-1861, 2611-2666[a]	0.00	3212-3473	0.56	0.67	0.22	0.00
Dioryctria sylvestrella	2	0.09	0.17	2471-3070, 3392-3472	0.00	1761-2466, 3675-3387	0.00	0.17	0.00	0.17
Dioryctria yiai	2	0.30	0.67	3073-2291	0.00	2701-2991	0.50	0.33	0.17	0.33
Feltia jaculifera	3	2.25	3.22	2602-2607	1.67	3136-3141	2.16	3.22	1.78	2.11
Hyles euphorbiae	6	0.20	0.39	1994-2204[b]	0.06	2594-2664	0.28	0.06	0.29	0.23
Hyles gallii	6	0.19	0.34	3233-3458	0.07	2011-2046	0.15	0.18	0.2	0.34
Hyles robertsi peplidis	3	0.03	0.11	3419-3473	0.00	1776-3396	0.00	0.00	0.00	0.11
Hyles sammutii	4	1.36	1.69	2155-2205	0.00	3015-3121	1.57	1.31	0.94	1.25
Hyles tithymali	12	0.12	0.22	1801-1816	0.00	2222-2492	0.20	0.09	0.12	0.11
Lambdina fiscellaria	2	1.81	3.00	1862-1992	0.83	2962-3170	2.99	1.67	0.83	1.33
Papilio demodocus	2	0.74	1.33	2497-2652	0.17	1767	0.5	1.33	0.83	0.50
Papilio demoleus	2	2.71	4.50	2031-2046[b]	1.50	3103-3113	3.33	3.33	1.83	1.67
Papilio grosesmithi	2	0.26	0.50	3471	0.00	2031-2476, 3093-3168	0.17	0.17	0.17	0.33
Papilio machaon	3	0.95	1.83	1986-2081[b]	0.11	2981-3097	1.50	0.78	0.11	0.56
Diptera										
Anopheles gambiae	3	0.46	0.67	2352-2527, 3307-3407	0.22	1892-1877, 2002-2162	0.33	0.56	0.33	0.56
Anopheles melas	2	0.78	1.30	3267-3272	0.50	2082-2107, 2437-2492, 2802-3132	1.0	0.67	0.67	1.00
Chrysomya megacephala	2	0.17	0.67	1962-2117[b]	0.00	2562-3472	0.50	0.00	0.00	0.00
Chrysomya rufifacies	2	0.78	1.33	2042[b]	0.33	2982-3170	1.00	0.83	0.50	0.67
MEAN (± 1 st. dev)		0.68 (±0.76)	1.12 (±1.15)		0.28 (±0.49)		0.84 (±0.96)	0.76 (±0.96)	0.46 (±0.54)	0.58 (±0.59)

[a] Region of maximum divergence encompasses midpoint *Dy*#1856, which corresponds to the region used by DNA barcoding

[b] Region of maximum divergence encompasses midpoint *Dy*#2042

[c] Source of speciemens and original references are listed in Appendix 1.

Table 3- 2: Interspecific nucleotide divergence (%) for 23 sister species pairs. Values for full 2.3 kb of COI-COII divergence (Total), maximum divergence (Max), and minimum divergence (Min) are shown. Midpoints are shown in *D. yakuba* numbers, and regions with multiple equal divergence values were combined when less than 100 bp apart. Divergence was also examined for representative 600 bp fragments for the four primer regions described in Table 1.

Sister Species Pairs[c]	Total%	Max%	Midpoint(s)	Min%	Midpoint(s)	BC	J-P	G-M	P-E
Lepidoptera									
Choristoneura α lineage x *C.* β lineage	2.23	2.82	2470-2480	1.59	3187-3212	2.42	2.41	1.89	2.15
C. fumiferana x *C.* α+β lineage	2.47	3.96	2575-2615[b]	1.52	3187-3212	1.91	3.96	2.28	2.42
Dioryctria abietella x *D. abietivorella*	3.87	5.17	2497-2632[b]	1.91	3145-3210	4.29	5.17	2.20	3.81
D. reniculelloides x *D. pseudotsugella*	1.72	2.37	1830-1934, 3382-3472	0.83	2941-3001	2.04	1.18	1.35	2.37
D. rubella x *D. sylvestrella*	1.27	2.28	2692-2702[b]	0.83	2182-2307	1.18	1.94	1.18	1.43
D. zimmermani x *D. cambiicola*	0.31	0.67	2321-2486	0.16	2901-3472	0.34	0.33	0.17	0.30
Hyles annei x *H. euphorbarium*	2.93	5.00	2671-2691[b]	1.35	3472	3.24	4.64	2.72	1.52
H. gallii x *H. nicaea*	1.63	2.54	2620-2645[a]	1.04	2014-2024	1.15	2.37	2.14	1.66
Papilio canadensis x *P. glaucus*	1.28	2.37	1891-1951	0.50	3167-3182	2.20	1.18	1.01	0.67
P. demodocus x *P. erithonioides*	4.11	6.53	2632-2642[a]	2.98	1767	3.32	6.07	4.46	4.38
P. erostratus x *P. anchicides*	2.80	3.76	2706[b]	1.86	2326-2356	3.24	3.41	2.89	2.54
P. machaon x *P. hospiton*	3.47	4.94	2066	2.49	3152-3157	4.05	3.29	2.77	3.58
P. m. oregonius x *P. m. hippocrates*	1.37	2.54	1986-2081	0.17	2981-3097, 3277	2.03	1.18	0.17	0.838
P. memnon x *P. rumanzovia*	3.81	4.64	3377-3422	2.89	3252-3277	3.76	4.11	4.11	4.29
P. thoas x *P. cresphontes*	5.90	7.53	2671-2676, 3192-3202[b]	5.00	1836-1882	5.17	6.62	6.80	6.62
P. troilus x *P. palamedes*	6.76	9.02	2636-2651[a]	5.00	2236-2246	5.89	8.46	6.25	7.90
P. zelicaon x *P. polyxenes*	2.67	3.58	1796-1861, 2641-2746[a]	1.35	2236-2241	3.41	3.41	2.72	2.72
Yponemeuta cagnagella x *Y. padella*	0.18	0.50	1846-2006	0.00	2446-2691, 3308-3473	0.50	0.00	0.17	0.00
Diptera									
Anopheles gambiae x *A. arabiensis*	0.52	1.00	3212-3232	0.28	1892-1927	0.33	0.45	0.84	0.50
A. melas x *A. quadriannulatis*	1.45	2.20	1852	0.84	2592-2617	2.03	0.84	0.18	1.35
Chrysomya megacephala x *C. bezziana*	4.11	6.98	2542-2574[b]	2.54	3122-3147	3.32	6.62	3.06	3.58
C. norrisi x *C. variceps*	5.21	7.35	2317-2347	3.41	3078-3195	6.98	7.17	3.584	3.94
C. rufifacies x *C. albiceps*	3.06	4.55	2542-2567[b]	2.28	1767-1772	2.89	4.38	2.98	3.50
MEAN	2.74	4.01		1.77		2.86	3.44	2.43	2.69
(±1 st.dev.)	(±1.74)	(±2.31)		(±1.40)		(±1.71)	(±2.42)	(±1.80)	(±1.96)

[a] Region of maximum divergence encompasses midpoint *Dy*#2662.
[b] Region of maximum divergence occurs within 100 bp of midpoint *Dy*#2662
[c] Source of specimens and original references are listed in Appendix 1.

Table 3- 3: Nucleotide divergence and relative percentages of total COI-COII divergence sampled by PCR fragments commonly used in insect systematics. Mean divergence and relative divergence are calculated from 23 sister species pairs.

PCR Fragment	Primer Names	Length (bp)	Mean Absolute Divergence (%)	Mean Relative Divergence (%) (± 1 St. Dev %)
LCO[a]-HCO[a]	LCO1490 (1514)[b] HCO2198 (2173)[b]	658	2.81 (±1.76)	111.88 (±44.45)
Jerry[c]-Pat[c]	C1-J-2183 TL2-N-3014	800	3.24 (±2.22)	114.44 (±33.75)
George[d]-Marilyn[c]	C1-J-2792 C2-N-3389	596	2.40 (±1.78)	87.09 (±26.42)
Pierre[c]-Eva[d]	C2-J-3138 TK-N-3775	643	2.65 (±2.05)	87.24 (±30.12)
Total COI-COII			2.74 (±1.74)	100 (±0)

[a]Folmer et al. 1994
[b]*D. yakuba* bp number at 3' end
[c]Simon et al. 1994
[d]Bogdanowicz et al. 1993

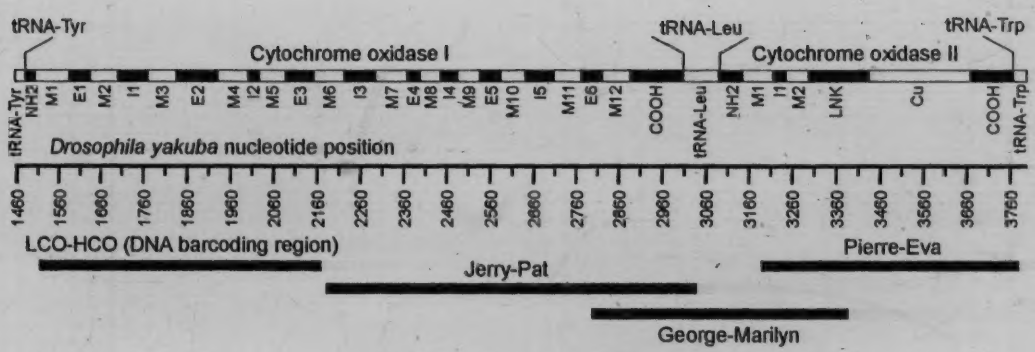

Figure 3- 1: Overview of COI-COII genes, with nucleotide locations as in *D. yakuba* (Clary and Wolstenholme 1985). Structural regions are shown along the upper bar and abbreviated according to Lunt et al. (1996). Commonly used COI-COII regions in insect systematics (Caterino et al. 2000) are identified by their forward and reverse primers (Bogdanowicz et al. 1993, Folmer et al. 1994, Simon et al. 1994).

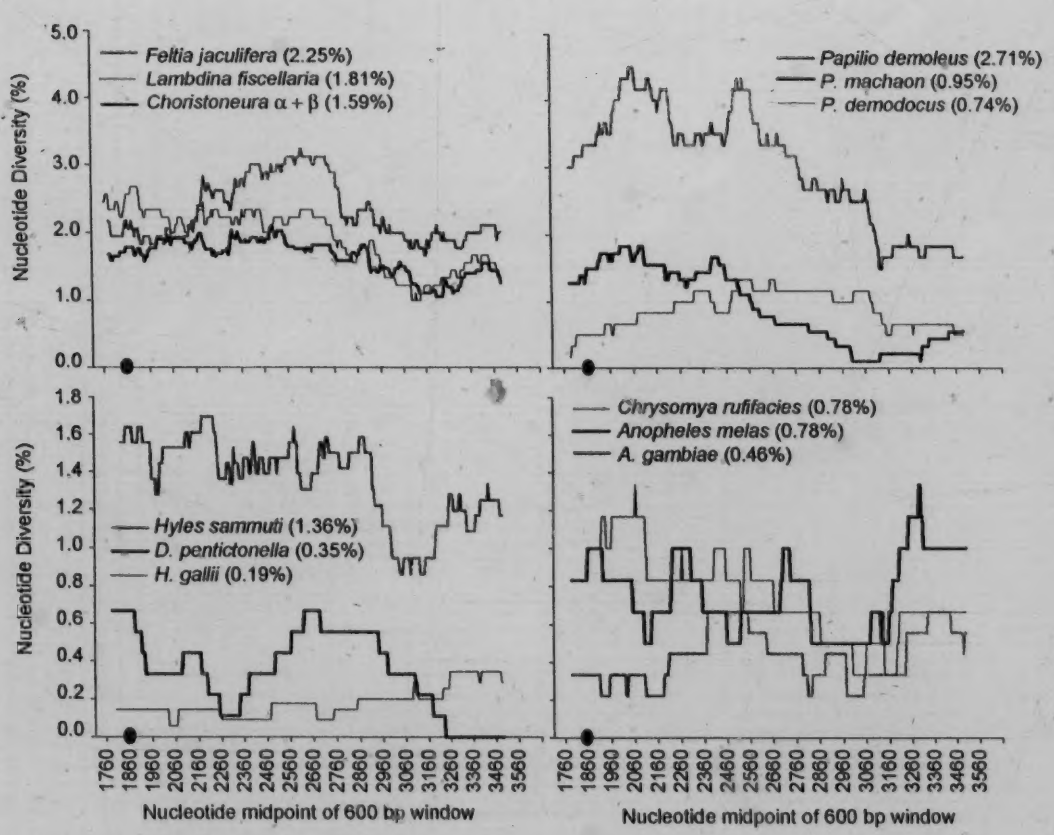

Figure 3- 2: Sliding window profiles of COI-COII (600 bp window; 5 bp step) of intraspecific nucleotide diversity for 12 of 23 species. Average nucleotide diversity for total COI-COII follows each species name. Black circles indicate DNA barcoding midpoint. Scales on Y-axes differ between upper and lower panels.

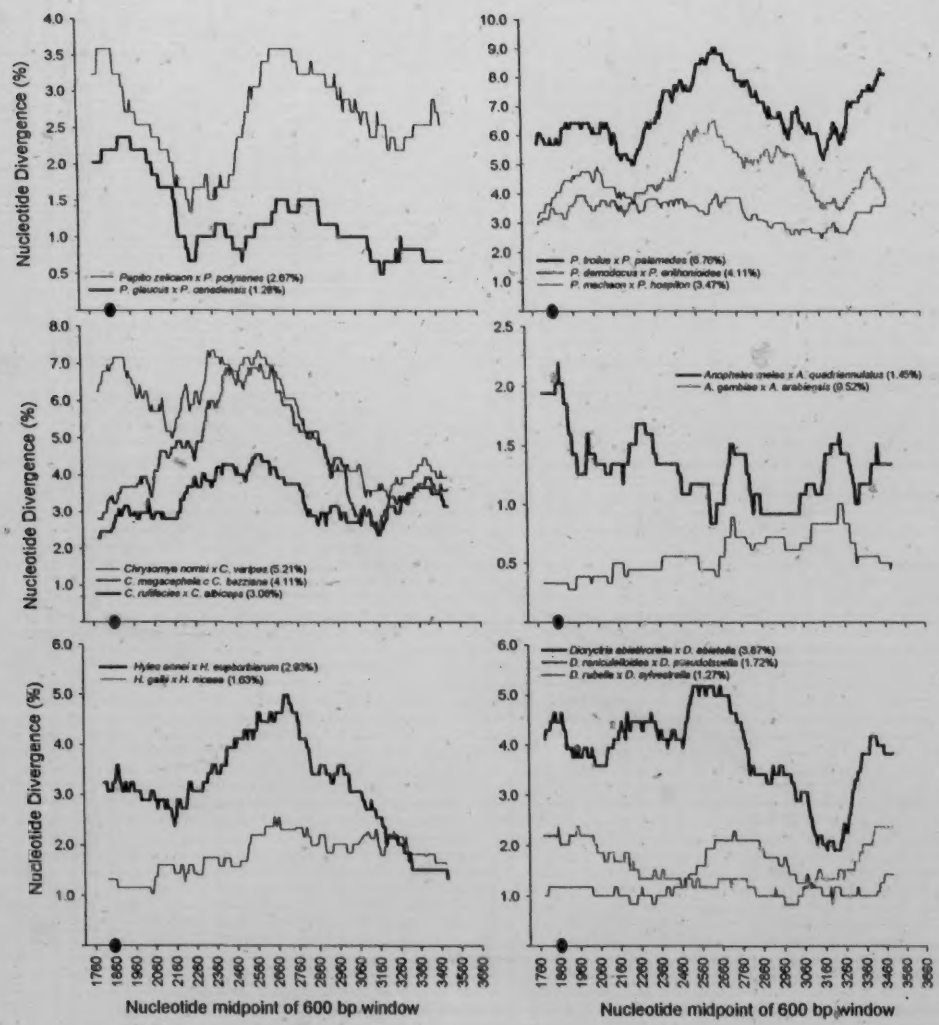
Nucleotide Divergence (%)
Papilio zelicaon x P. polyxenes (2.67%)
P. glaucus x P. canadensis (1.28%)
P. troilus x P. palamedes (0.70%)
P. demodocus x P. erithonioides (4.11%)
P. machaon x P. hospiton (3.47%)
Chrysomya nonisi x C. varipes (5.21%)
C. megacephala x C. bezziana (4.11%)
C. rufifacies x C. albiceps (3.06%)
Anopheles melas x A. quadriannulatus (1.45%)
A. gambiae x A. arabiensis (0.52%)
Hyles annei x H. euphorbiarum (2.93%)
H. gallii x H. nicaea (1.63%)
Diacrisia obliterrorella x D. obliqua (3.67%)
D. ranculelloides x D. pseudobrunnea (1.72%)
D. rubella x D. sylvestrella (1.27%)
Nucleotide midpoint of 600 bp window

Figure 3- 3: Sliding window profiles of COI-COII (600 bp window; 5 bp step) of nucleotide divergence for 15 of 23 sister species pairs. Total COI-COII nucleotide divergence follows each species name. Black circles indicate DNA barcoding midpoint. Scales on Y-axes are variable.

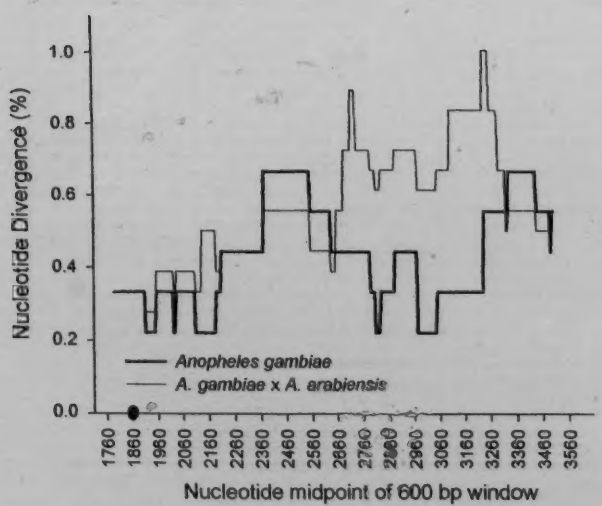

Figure 3- 4: Sliding window profiles of COI-COII (600 bp window; 5 bp step) of intraspecific nucleotide diversity (*A. gambiae*) and interspecific divergence (*A. gambiae* x *A. arabiensis*) showing regions of overlap. Black circle indicates DNA barcoding midpoint.

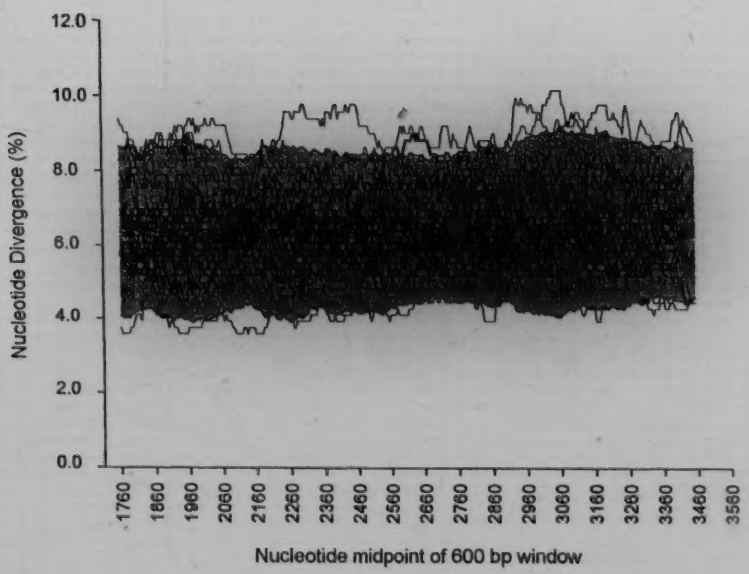

Figure 3- 5: Sliding window profile of *P. troilus* x *P. palamedes* and profiles of 50 randomly generated sequence pairs. Original profile (thick bold line), mean random divergence (thin bold line) and ± 95% CI (shaded region) of random profiles are shown.

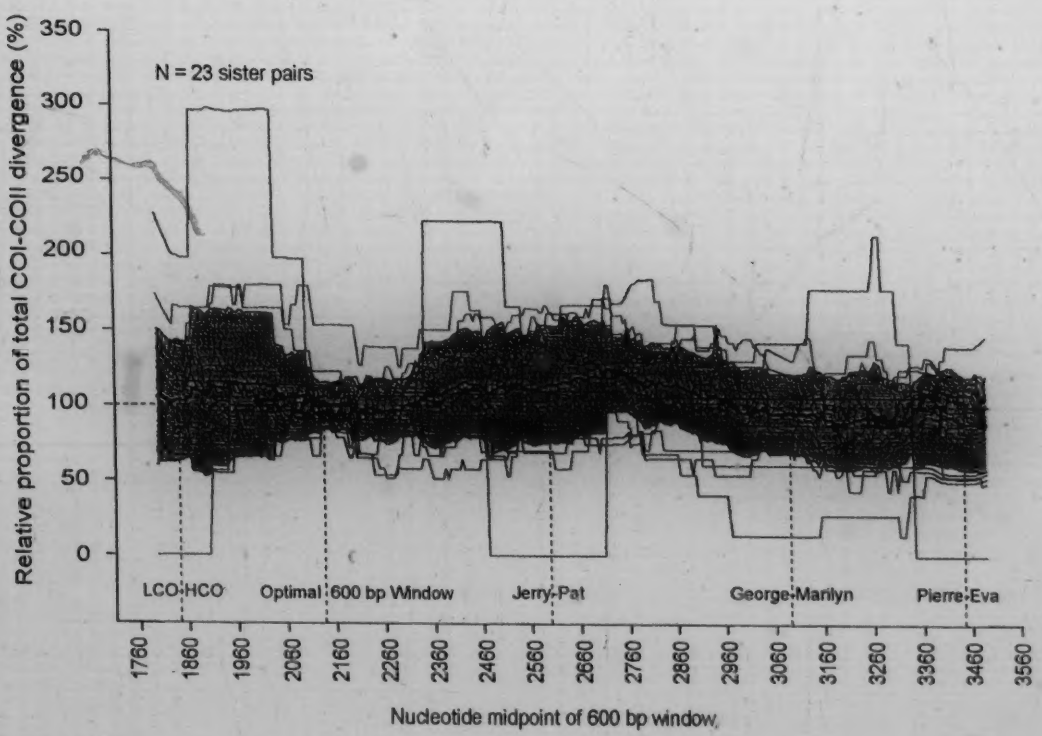

Figure 3- 6: Sliding window profile of mean relative divergence (bold) and ± 1 standard deviation (shaded) for all 23 sister species pairs. Individual profiles (grey lines) show individual sister species pairs. Midpoint locations of commonly used COI-COII PCR regions and the "optimal" 600 bp region are shown.

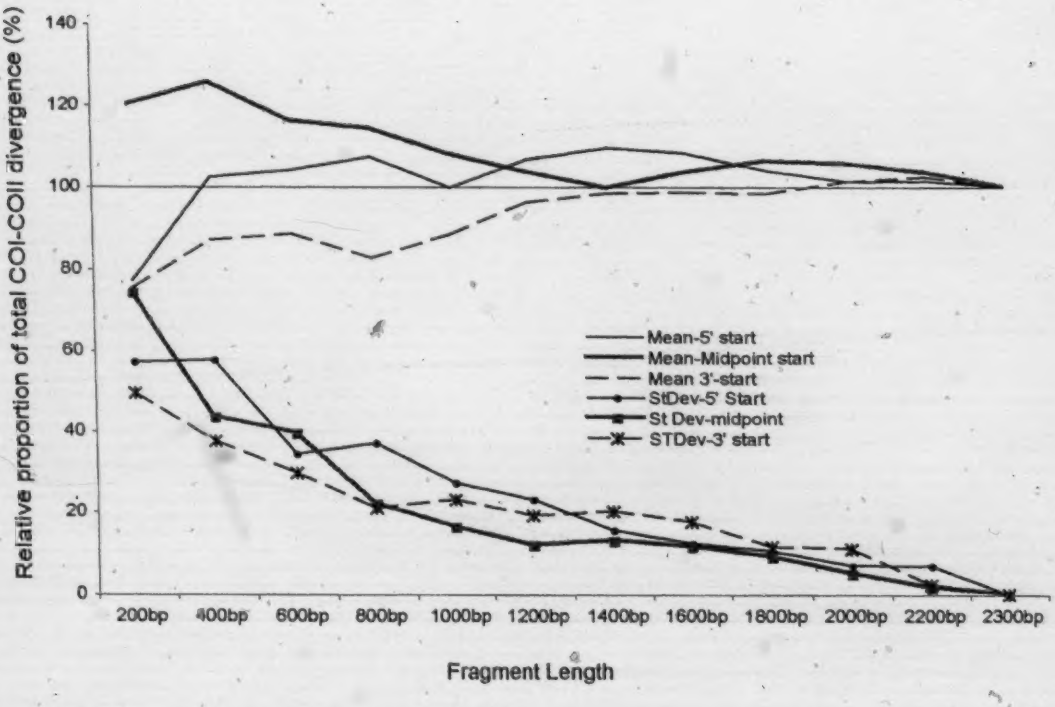

Figure 3- 7: Mean relative nucleotide divergence and variance (1 standard deviation) for fragments of increasing length starting at either the 5' end of COI, the midpoint of COI-COII, or the 3' end of COII for 23 sister species. Fragment length started at 200 bp, and increased in 200 bp increments until the entire sequence was included

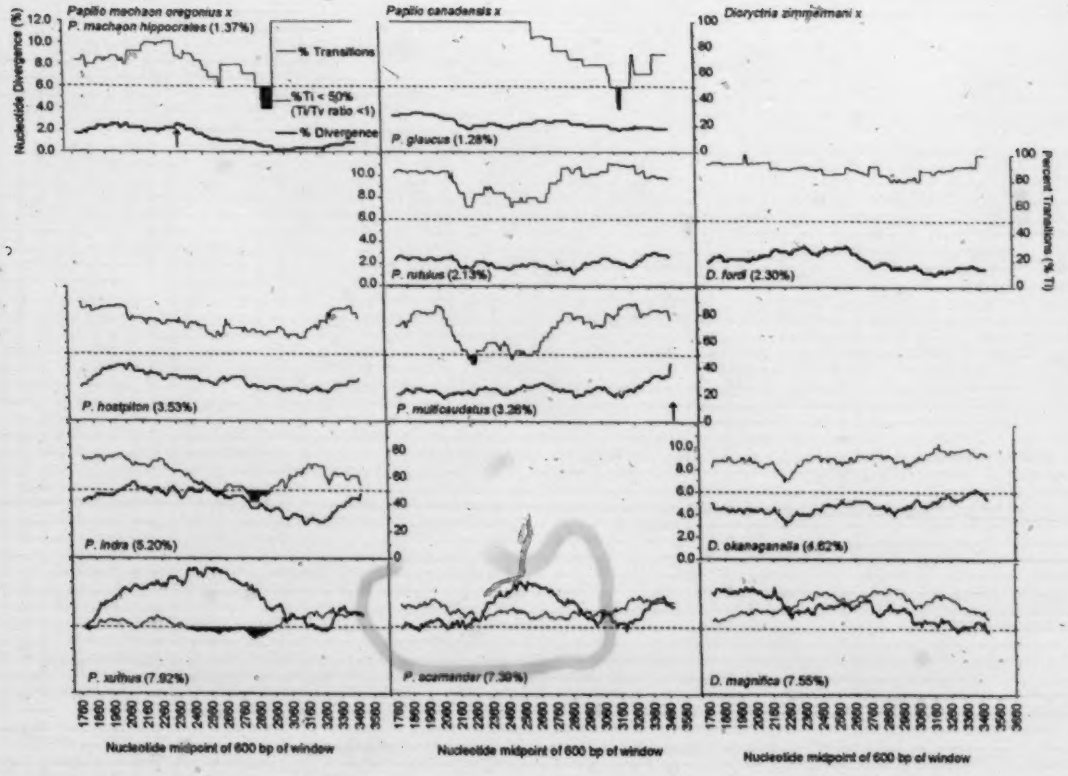

Figure 3- 8: Phylogenetic time series for *Papilio machaon oregonius*, *Papilio canadensis*, and *Dioryctria zimmermani* with phylogenetic series arranged in columns and similar total divergences arranged in rows. Each species is compared to increasingly more diverged congeners of roughly equal increments (where available), with total divergence indicated for each comparison. Sliding window profiles of COI-COII nucleotide divergence (bold) and percent transitions (fine) are presented for each species pair. Left axes indicate percent nucleotide divergence, and right axes represent percent transitions. Areas of < 50% transitions (dotted line) are shaded to indicate regions of saturation. Arrows indicate regions of relatively high divergence and high percent transitions.

Literature Cited

Armstrong, K. F., and S. L. Ball. 2005. DNA barcodes for biosecurity: invasive species identification. Philisophical Transactions of the Royal Society of London Series B 360: 1813-1823.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18: 489-522.

Baker, C. S., and S. R. Palumbi. 1994. Which whales are hunted? A molecular genetic approach to monitoring whaling. Science 265: 1538-1539.

Barrett, R. D. H., and P. D. N. Hebert. 2005. Identifying spiders through DNA barcodes. Canadian Journal of Zoology 83: 481-491.

Bartlett, S. E., and W. S. Davidson. 1992. FINS (Forensically Informative Nucleotide Sequencing): a procedure for identifying animal origin of biological specimens. BioTechniques 12: 408-411.

Bartolome, C., X. Maside, S. Yi, A. L. Grant, and B. Charlesworth. 2005. Patterns of selection on synonymous and nonsynonymous variants in *Drosophila mirandai*. Genetics 169: 1495-1507.

Beard, C. B., D. M. Hamm, and F. H. Collins. 1993. The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. Insect Molecular Biology 2: 103-124.

Besansky, N. J., J. Krzywinski, T. Lehmann, F. Simard, M. Kern, O. Mukabayire, D. Fontenill, Y. Touré, and N. F. Sagnon. 2003. Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. Proceedings of the National Academy of Sciences of the United States of America 100: 10818-10823.

Blair, C. P., W. G. Abrahamson, J. A. Jackman, and L. Tyrrell. 2005. Cryptic speciation and host-race formation in a purportedly generalist tumbling flower beetle. Evolution 59: 304-316.

Blair, J. E., and S. B. Hedges. 2005. Molecular phylogeny and divergence times of deuterostome animals. Molecular Biology and Evolution 22: 2275-2284.

Bogdanowicz, S. M., W. E. Wallner, J. Bell, T. M. O' Dell, and R. G. Harrison. 1993. Asian gypsy moths (Lepidoptera, Lymantridae) in North America - Evidence from molecular data. Annals of the Entomological Society of America 86: 710-715.

Caterino, M. S., S. Cho, and F. A. H. Sperling. 2000. The current state of insect molecular systematics: A thriving tower of Babel. Annual Review of Entomology 45: 1-54.

Caterino, M. S., and F. A. H. Sperling. 1999. *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. Molecular Phylogenetics and Evolution 11: 122-137.

Clary, D. O., and D. R. Wolstenholme. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. Journal of Molecular Evolution 22: 252-271.

Dayrat, B. 2005. Towards integrative taxonomy. Biological Journal of the Linnean Society 85: 407-415.

Du, Y., A. D. Roe, and F. A. H. Sperling. 2005. Phylogenetic framework for *Dioryctria* (Lepidoptera: Pyralidae: Phycitinae) based on combined analysis of mitochondrial DNA and morphology. Canadian Entomologist 137: 685-711.

Dunn, C. P. 2003. Keeping taxonomy based in morphology. Trends in Ecology & Evolution 18: 270-271.

Ebach, M. C., and C. Holdrege. 2005. DNA barcoding is no substitute for taxonomy. Nature 434: 697.

Edwards, S. V., W. B. Jennings, and A. M. Shedlock. 2005. Phylogenetics of modern birds in the era of genomics. Proceedings of the Royal Society of London Series B 272: 979-992.

Erpenbeck, D., J. N. A. Hooper, and G. Wörheide. 2005. CO1 phylogenies in diploblasts and the 'Barcoding of Life' - are we sequencing a suboptimal partition? Molecular Ecology Notes: 1-4.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294-299.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution and Systematics 34: 397-423.

Galtier, N., D. Enard, Y. Ranondy, E. Bazin, and K. Belkhir. 2006. Mutation hot spots in mammalian mitochondrial DNA. Genome Research 16: 215-222.

Goetze, E. 2003. Cryptic speciation on the high seas; global phylogenetics of the copepod family Eucalanidae. Proceedings of the Royal Society of London Series B 270: 2321-2331.

Gunning, K. M., E. McIntyre, C. Brown, K. Chansky, J. Bourey, B. Johnson, A. Wheaton, and C. Wike. 2002. ABI PRISM® 3100 Genetic Analyzer: Further advances to expand productivity. Applied Biosystems.

Hanner, R. 2005. Proposed standards for BARCODE records in INSDC (BRIs), pp. 3. National Museum of Natural History, Washington D.C.

Hebert, P. D. N., and T. R. Gregory. 2005. The promise of DNA barcoding for taxonomy. Systematic Biology 54: 852-859.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003a. Biological Identification through DNA barcodes. Proceedings of the Royal Society of London Series B 270: 313-321.

Hebert, P. D. N., S. Ratnasingham, and J. R. deWaard. 2003b. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London Series B (Suppl.) 270: S96-S99.

Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemiak, and C. M. Francis. 2004. Identification of birds through DNA barcodes. Public Library of Science Biology 2: e312.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2005. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the United States of America 101: 14812-14817.

Hendrixson, B. E., and J. E. Bond. 2005. Testing species boundaries in the *Anthrodiaetus unicolour* complex (Araneae: Mygalomorphae: Antrodiaetidae): "Paraphyly" and cryptic diversity. Molecular Phylogenetics and Evolution 36: 405-416.

Honeycutt, R. L., M. a. Nedbal, R. M. Adkins, and L. L. Janecek. 1995. Mammalian mitochondrial-DNA evolution - a comparison of the cytochrome-b and cytochrome-c oxidase-II genes. Journal of Molecular Evolution 40: 260-272.

Hundsoerfer, A. K., I. J. Kitching, and M. Wink. 2005. A molecular phylogeny of the hawkmoth genus *Hyles* (Lepidoptera: Sphingidae: Macroglossinae). Molecular Phylogenetics and Evolution 35: 442-458.

Innan, H., and M. Nordborg. 2002. Recombination or mutational hot spots in human mtDNA? Molecular Biology and Evolution 19: 1122-1127.

Johnson, K. P., B. L. Williams, D. M. Drown, R. J. Adams, and D. H. Clayton. 2002. The population genetics of host specificity: genetic differentiation in dove lice (Insects: Phthiraptera). Molecular Ecology 11: 25-38.

Lambert, D. M., A. Baker, L. Huynen, O. Haddrath, P. D. N. Hebert, and C. D. Millar. 2005. Is a large-scale DNA-based inventory of ancient life possible? Journal of Heredity 96: 279-284.

Lin, C.-P., and B. N. Danforth. 2003. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Molecular Phylogenetics and Evolution 30: 686-702.

Lipscomb, D., N. Platnick, and Q. Wheeler. 2003. The intellectual content of taxonomy: a comment on DNA taxonomy. Trends in Ecology and Evolution 18: 65-66.

Lunt, D. H., D.-X. Zhang, J. M. Szymura, and G. M. Hewitt. 1996. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. Insect Molecular Biology 5: 153-165.

Maddison, W. P., and D. R. Maddison 2004. Mesquite: A modular system for evolutionary analysis computer program, version 1.05. By Maddison, W. P., and D. R. Maddison.

Meyer, C. P., and G. Paulay. 2005. DNA barcoding: error rates based on comprehensive sampling. Public Library of Science, Biology 3: e422.

Meyer, A., and R. Zardoya. 2003. Recent advances in the (molecular) phylogeny of vertebrates. Annual Review of Ecology, Evolution and Systematics 34: 311-338.

Monaghan, M. T., M. Balke, J. Pons, and A. P. Vogler. 2006. Beyond barcodes: complex DNA taxonomy of a South Pacific island radiation. Proceedings of the Royal Society B-Biological Sciences 273: 887-893.

Moritz, C., and C. Cicero. 2004. DNA barcoding: promise and pitfalls. Public Library of Science Biology 2: e354.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Olson, R. R. 1991. Whose larvae? Nature 351: 352-353.

Page, T. J., S. C. Choy, and J. M. Hughes. 2005. The taxonomic feedback loop: symbiosis of morphology and molecules. Biology Letters 1: 139-142.

Penton, E. H., P. D. N. Hebert, and T. J. Crease. 2004. Mitochondrial DNA variation in North American populations of *Daphnia obtusa*: continentalism or cryptic endemism? Molecular Ecology 13: 97-107.

Posada D. and Crandall K.A. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14 (9): 817-818

Regier, J. C., J. W. Shultz, and R. E. Kambic. 2005. Pancrustacean phylogeny: hexapods are terrestrial crustaceans and maxillopods are not

monophyletic. Proceedings of the Royal Society B-Biological Sciences 272: 395-401.

Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.

Saraste, M. 1990. Structrual features of cytochrome oxidase. Quarterly Review of Biophysics 23: 331-366.

Scheffer, S. J., R. M. Giblin-Davis, G. S. Taylor, K. A. Davies, M. Purcell, M. I. Lewis, J. Goolsby, and T. D. Center. 2004. Phylogenetic relationships, species limits, and host specificity of gall-forming *Fergusonina* flies (Diptera: Fergusoninidae) feeding on *Melaleuca* (Myrtaceae). Annals of the Entomological Society of America 97: 1216-1221.

Seberg, O., C. J. Humphries, S. Knapp, D. W. Stevenson, G. Petersen, N. Scharff, and N. M. Andersen. 2003. Shortcuts in systematics? A commentary on DNA-based taxonomy. Trends in Ecology and Evolution 18: 63-65.

Sharley, D. J., V. Pettigrove, and Y. M. Parsons. 2004. Molecular identification of *Chironomus* spp. (Diptera) for biomonitoring of aquatic ecosystems. Australian Journal of Entomology 43: 359-365.

Sheldon, F. H., and A. H. Bledsoe. 1993. Avian Molecular Systematics, 1970S to 1990S. Annual Review of Ecology and Systematics 24: 243-278.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87: 651-701.

Sperling, F. A. H. 2003. DNA barcoding: *Deus ex Machina*. Newletter of the Biological Survey of Canada (Terrestrial Arthropods). Opinion Page [online] 22 (1). Available from http://www.biology.ulaberta.ca/bsc/news22_2/opinionpage.htm [cited 15 January 2006].

Sperling, F. A. H., and D. A. Hickey. 1994. Mitochondrial DNA sequence variation in the spruce budworm species complex (*Choristoneura*: Lepidoptera). Molecular Biology and Evolution 11: 656-665.

Sperling, F. A. H., J.-F. Landry, and D. A. Hickey. 1995. DNA-based identification of introduced ermine moth species in North America (Lepidoptera: Yponomeutidae). Annals of the Entomological Society of America 88: 155-162.

Sperling, F. A. H., R. Byers, and D. A. Hickey. 1996. Mitochondrial DNA sequence variation among pheromotypes of the dingy cutworm, *Feltia jaculifera* (Gn.) (Lepidoptera: Noctuidae). Canadian Journal of Zoology 74: 2109-2117.

Sperling, F. A. H., A. G. Raske, and I. S. Otvos. 1999. Mitochondrial DNA sequence variation among populations and host races of *Lambdina fiscellaria* (Gn.) (Lepidoptera: Geometridae). Insect Molecular Biology 8: 97-106.

Stoneking, M. 2000. Hypervariable sites in the mtDNA control region are mutational hotspots. American Journal of Human Genetics 67: 1029-1032.

Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference, pp. 407-514. *In* D. M. Hillis, C. Moritz and B. Mable [eds.], Molecular Systematics, 2nd ed. Sinauer Associates, Sunderland, Massachusetts.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105: 437-460.

Tan, S. H., E. Mohd-Aris, B. Omar, J. Surin, and Z. Mohamed. 2005. DNA-based characteristics of forensically important insect species in Malaysia. GenBank.

Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, and A. P. Volger. 2003. A plea for DNA taxonomy. Trends in Ecology and Evolution 18: 70-74.

Wells, J. D., and F. A. H. Sperling. 1999. Molecular phylogeny of *Chrysomya albiceps* and *C. rufifacies* (Diptera: Calliphoridae). Journal of Medical Entomology 36: 222-226.

Wells, J. D., and F. A. H. Sperling. 2001. DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). Forensic Science International 3065: 1-6.

Wells, J. D., F. Introna, G. Di Vella, C. P. Campobasso, J. Hayes, and F. A. H. Sperling. 2001. Human and insect mitochondrial DNA analysis from maggots. Journal of Forensic Sciences 46: 685-687.

Whiteman, N. K., D. Santiago-Alarcon, K. P. Johnson, and P. G. Parker. 2004. Differences in straggling rates between two genera of dove lice (Insecta: Phthiraptera) reinforce population genetic and cophylogenetic patterns. International Journal of Parasitology 34: 1113-1119.

Will, K. W., B. D. Mishler, and Q. D. Wheeler. 2005. The perils of DNA barcoding and need for integrative taxonomy. Systematic Biology 54: 844-851.

Will, K. W., and D. Rubinoff. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20: 47-55.

Zakharov, E. V., M. S. Caterino, and F. A. H. Sperling. 2004. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). Systematic Biology 53: 193-215.

Zhang, D.-X., and G. M. Hewitt. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular Ecology 12: 563-58

Chapter 4: Delimitation of two sympatric *Dioryctria* species using an integrative taxonomic approach.

Introduction

Species delimitation is part of the process of identifying and delineating distinct organismal entities in nature. Many studies rely on species as their primary unit, and inaccurate delimitations could seriously affect biodiversity assessments, invasive species recognition and biological control projects. Currently, the methods used to delimit and classify species are undergoing a revolution (Sites and Marshall 2003, Dayrat 2005). Society is faced with a taxonomic crisis (Godfray 2002) with the continual loss of taxonomists and taxonomic knowledge, and as a result, traditional morphological taxonomy cannot keep pace with the need for timely species delimitations (Godfray 2002, Mallet and Willmot 2003, Wilson 2003). In parallel with these concerns, the concept of integrative taxonomy has become increasingly popular in taxonomic circles (Dayrat 2005, Will et al. 2005). Rather than concentrating on a single character type (e.g. morphological variation or DNA sequence variation), taxonomists are now integrating a wide range of characters, including molecular data, morphology, behavior, and geography, to assess species boundaries and delimit species.

DNA sequences are becoming increasingly popular as a means to delimit species. These data are generally numerous, easy to obtain, and bridge the gap between intra- and interspecific variation (Templeton 2001). MtDNA genes are particularly popular for a range of species-level questions as they evolve rapidly, are maternally inherited and effectively haploid, have limited recombination, and are robust against degradation (Avise et al. 1987). The use of mtDNA for species delimitations has seen a rapid increase as projects such as DNA barcoding (Hebert et al. 2003, Hebert and Gregory 2005) become more common. This method uses a short mtDNA fragment from the cytochrome c oxidase (COI) gene and identifies species or detects potential cryptic species based on relative genetic divergence (Hebert et al. 2003, Hebert et al. 2004, Hebert et al. 2005). Although these methods can be very effective, relying solely on these data can result in both over- and under-estimations of species diversity (Will and

Rubinoff 2004, Meyer and Paulay 2005). MtDNA is also prone to incomplete lineage sorting, retained ancestral polymorphisms, and introgression due to hybridization, which could lead to incorrect species delimitation (Funk and Omland 2003).

It is desirable, therefore, to use several independent characters, such as multiple molecular loci when trying to delimit species (Sperling 2003, Dayrat 2005). In addition, other sources of data, such as morphology, host preference, geographic range and pheromones should be used when identifying species boundaries (Coyne and Orr 2004), as suggested by supporters of integrative taxonomy.

In order to assess the accuracy of various types of characters, it is also useful to state explicitly which species definition is being employed. Discussions surrounding species concepts have continued to be contentious, and numerous definitions have been described (Coyne and Orr 2004, see their Appendix A). For the purpose of this study I consider a species to be a cohesive group of populations that maintains its genomic integrity with other such groups, even in the presence of gene flow (Sperling 2003). Allopatric populations may be considered separate species if they have levels of genetic divergence similar to that of distinct, sympatric species.

Species in the genus *Dioryctria* Zeller provide a test case for application of an integrative taxonomic approach via assessment of several independent character sets. The genus contains several species complexes where species are highly variable, broadly sympatric, and are currently diagnosed by minor morphological differences. Nearctic members of the *schuetzeella* group Mutuura and Munroe, *Dioryctria reniculelloides* Mutuura and Munroe and *D. pseudotsugella* Munroe, are typical examples of the problems found within this genus. *Dioryctria reniculelloides* is a primarily boreal species, and feeds on needles on primarily *Picea* sp. Additional hot records include *Abies* sp., *Pseudotsuga menziesii,* and *Tsuga heterophylla. Dioryctria pseudotsugella* is found throughout the west, from southwestern Canada to Arizona. This species also feeds on needles, although it is found primarily on *Pseudotsuga menziesii* and *Abies* sp. Based on their original descriptions, *D. reniculelloides* and *D. pseudotsugella* were separated primarily by geographic distribution, larval host plant, and minor forewing characteristics, such as length, overall colour, and the colour and size of a subbasal scale patch (Munroe 1959, Mutuura and Munroe 1973, Neunzig 2003). However, these two

putative taxa are broadly sympatric throughout southern Alberta, British Columbia and the western United States, share larval host records, and intermediate morphological forms are common (Sopow et al. 1996). To accurately delimit these taxa and identify species boundaries, a detailed examination of morphometric and molecular data are needed to clarify this species problem (Sopow et al. 1996).

The objective of this study is to quantify the genetic, morphological and ecological differences in specimens belonging to the *schuetzeella* group, and test the traditional delimitation of two putatively distinct species using multiple independent lines of evidence. I examine three independent molecular markers, four morphometric forewing traits, and larval host plant use to address the following questions: 1) Does the level of gene flow and genetic differentiation among populations support the traditional taxonomic hypothesis of two distinct species?; 2) Do patterns of morphological variation correlate with genetic variation or larval host plant use?; 3) How congruent are the character sets used to evaluate species boundaries?; and 4) Are there diagnostic morphological, molecular or behavioural characters that can improve identification of the delimited species?

Materials and Methods

Species and Study Area.

Specimens of the *schuetzeella* group were obtained from localities across Canada and the western United States (Fig. 4-1). Collecting was concentrated in southern Alberta and British Columbia, where the two putative taxa are believed to overlap (Neunzig 2003). A total of 108 specimens were either collected as adults using UV light traps, or as larvae and then reared on the foliage of their host plant (Appendix 2). Specimens obtained for molecular analysis were either live, or preserved in 100% EtOH, then frozen at -70°C. Vouchers and specimen images were deposited in the E.H. Strickland Entomological Museum, University of Alberta. COI sequences for a number of specimens were available on GenBank from previous studies and have been incorporated into this study: four specimens from Corvallis, Oregon (accession nos. DQ296160 (2 specimens with identical haplotypes), DQ296161, and DQ296162), one

specimen from Pritchard, British Columbia (DQ296164), and one specimen from Adams Lake, British Columbia (DQ296163) (all from Roe et al. in press). As well, COI sequence for a single specimen of *D. schuetzeella* Heinemann, a European member of the same species group, was available from GenBank (accession no. AJ868570) (Knolke et al. 2005) and included in this study for in-group comparison. Two *Dioryctria* species in the *auranticella* group Mutuura and Munroe (*D. auranticella* (Grote), accession no. DQ296157 and *D. rossi* Munroe, DQ296159) (Roe et al. in press) were included as outgroups for phylogenetic analyses. An additional 177 specimens were obtained for morphological examinations from entomological collections throughout Canada and the United States (Fig. 4-1; Appendix 2).

Laboratory Methods

Total genomic DNA was extracted from two legs or the whole thorax of 96 specimens collected in Alberta, British Columbia, Alaska and Oregon (Fig. 4-1; Table 4-1) following procedures outlined in Roe et al. (in press).

Three independent molecular makers were examined during the course of this study, one from mtDNA and two from nuclear DNA. The mtDNA marker was a 475 bp region of COI, corresponding to base pairs 2184-2658 of the *Drosophila yakuba* sequence (Clary and Wolstenholme 1985) and was sequenced from each individual. Previously published primers Jerry (C1-J-2183 5'-CAACATTTATTTTGATTTTTTGG-3') (Simon et al. 1994) and MilaIII (C1-N-2659 5'-ACTAATCCTGTGAATAAAGG -3') (Laffin et al. 2004) were used to amplify the region of COI. New COI sequences were deposited in GenBank under accession numbers DQ630944-DQ630959). PCR amplification, purification, cycle sequencing and product visualization protocols are outlined in Roe et al. (in press).

The first of two nuclear markers examined during the course of this study was the second internal transcribed spacer unit (ITS2), and sequences were either 491 or 499 bp in length due to the presence of several indels. A total of 29 individuals from 13 localities were sequenced for ITS2, with at least one representative from each mtDNA haplotype group. A new forward primer ITS3b (5'-GGGTCGATGAAGAACGCAST-3') and a previously designed reverse primer ITS4 (5'-TCCTCCGCTTATTGATATGC-3') (White

et al. 1990) were used to amplify ITS2. New ITS2 sequences were deposited in GenBank under accession numbers DQ792571-DQ792575. PCR amplifications were performed in a 50 μL reaction using 4.0 μL extracted DNA, 2 μL each of the two heterologous primers in 5 pmol/μL concentrations, 1 μL of 10 mmol/μL dNTPs (Roche Diagnostics, Indianapolis, Indiana); 5 μL 10x PCR reaction buffer containing 15 mmol/μL $MgCl_2$ (Promega Corporation, Madison, Wisconsin); 3 μL of 25 mmol/L $MgCl_2$ (Promega), 0.5μL Taq-polymerase (approximately 5 U/μL, Pickard Lab, University of Alberta, Edmonton, Alberta) and double-distilled H_2O (Millipore Corporation, Billerica, MA) to make up the remaining reaction volume. The PCR profile used to amplify ITS2 was as follows: 94°C for 2 min initial denaturation; 94° 30 sec denaturing, 55°C 1 min annealing, 72° 2 min extension, cycled 34 times; and 72°C 10 min final extension. PCR purification, cycle sequencing and product visualization follows Roe et al. (in press).

The second nuclear marker examined was a 365 bp fragment of elongation factor 1-α (EF1-α), corresponding to base pairs 2338 - 2702 of *Drosophila melanogaster* Meigen sequence (Hovemann et al. 1988). Previously developed forward primer 'Cho' (E234F: 5'-GTCACCATCATYGACGC-3') and reverse primer 'Juke' (E600rc: 5'-CTCCTTACGCTCAACATTC-3') (Reed and Sperling 1999) were used to amplify the EF1α fragment. New EF1α sequences were deposited in GenBank under accession numbers.DQ792594-DQ792598. PCR amplifications were performed in a 50μL reaction using 3.0 - 5.0 μL extracted DNA, 2 μL each of the two heterologous primers in 5 pmol/μL concentrations, 1 μL of 10 mmol/μL dNTPs (Roche Diagnostics); 5 μL 10x PCR reaction buffer containing 15 mmol/μL $MgCl_2$ (Promega); 0.5 μL of 25 mmol/L $MgCl_2$ (Promega), 0.5μL Taq-polymerase (approximately 5 U/μL, Pickard Lab) and double-distilled H_2O (Millipore Corporation, Billerica, MA) to make up the remaining reaction volume. The PCR profile used to amplify ITS2 was as follows: 94°C for 2 min initial denaturation; 94° 30 sec denaturing, 55°C 1 min annealing, 72° 2 min extension, cycled 34 times; and 72°C 10 min final extension. PCR purification, cycle sequencing and product visualization follows (Roe et al. in press).

Genetic Analyses

Identical sequences were combined into unique haplotypes (COI) or genotypes (ITS2 and EF1α). All EF1α sequences were homozygous, as judged on the basis of a lack of double peaks in chromatograms from both directions, while a single site was heterozygous in ITS2 with double peaks consistently evident in complementary strands, so genotypes were scored to reflect this variability (Tables 4-2 and 4-4).

Phylogenetic reconstructions were only conducted with mtDNA haplotypes, due to the low numbers of genotypes in ITS2 and EF1α. Relationships among mtDNA haplotypes were estimated using maximum parsimony (MP) in PAUP* version 4.0b10 (Altivec) (Swofford 2002). MP analysis was performed using a heuristic search with the following parameters: stepwise addition, 100 random addition replicates, and tree bisection reconnection (TBR), with all characters unordered and equally weighted. Multiple trees were presented as a strict consensus, and branch support was obtained using bootstap support values. Bootstrap support was obtained using 100 bootstrap replicates and the same heuristic search methods described above.

Estimation of genetic differentiation χ^2 (Nei 1987, Hudson et al. 1992) or Nst (Lynch and Crease 1990), nucleotide diversity π (Nei 1987, equation 10.5), and haplotype diversity h (Nei 1987, equation 8.4) were obtained using DNAsp (Rozas et al. 2003). SAMOVA (Spatial Analysis of Molecular Variance) (Dupanloup et al. 2002) was used to estimate gene flow (F_{ST}) among population groupings. Optimal numbers of groups were chosen by maximizing the among group variance, while minimizing variation between locality samples within groups.

A haplotype network was constructed using TCS vers. 1.21 (Clement et al. 2000) for all COI sequences, but was not attempted with either ITS2 or EF1α sequences due to lack of geographic variation and reduced sampling. Designation of nested clades was based on previously published guidelines (Templeton and Sing 1993, Templeton et al. 1995). A nested contingency analysis was implemented in GeoDis ver. 2.4 (Posada et al. 2000) to determine if there was significant genetic and geographic structuring in the previously defined clades of the haplotype network. Results of this analysis were interpreted using the November 11, 2005 revised inference key developed from Templeton et al. (1995), which is available on the GeoDis website (http://darwin.uvigo.es/software/geodis.html).

Morphometric Measurements

All specimens were photographed at 12X magnification using a Nikon CoolPix 990 digital camera mounted on a Wild dissecting scope (Model: M5-51289), using two separate fiber optic light sources. Camera settings were as follows: exposure "Aperture", image size "Fine"; f-stop 2.8; ISO 100; manual focus; white balance set manually using a white background. Forewing images of all specimens measured for morphometric characters were deposited in E.H. Strickland Entomological Museum, University of Alberta. All measurements were obtained using ImageJ vers. 1.36b (Rasband 2006; http://rsb.info.nih.gov/ij/). Four forewing characters were scored to quantify morphological variation among specimens. Measurements were chosen to correspond with diagnostic forewing characters used in the original species descriptions (Munroe 1959, Mutuura and Munroe 1973, Neunzig 2003). Maximum forewing length (Length), ground colour of the entire forewing (ADS), area of pinkish-orange subbasal patch (ASBA), and modal grayscale scale colour in the subbasal patch (Mode) were measured for each individual (Fig. 4-2). Maximum forewing length was measured from the base of the wing beside the body wall to the tip of the forewing, not including the fringe. ADS was measured as the area of dark scales on the overall surface of the forewing. To obtain this measurement, wing images were imported into Adobe Photoshop vers. 7.0.1 (Adobe Systems Inc. San Jose, CA), converted to grayscale (8-bit), reduced to 800 x 600 pixels, and Thresholds were adjusted to 128. The adjusted images were imported into ImageJ, thresholds were adjusted to 0 and 128, and the area of black in each forewing was recorded in square millimeters. ABSA was recorded in square millimeters, and the modal greyscale pixel colour of the patch was also recorded. A total of 285 specimens (including specimens used for molecular analyses) were measured, although only 195 specimens were measured for all four morphometric characters, due to damaged wings or large areas of missing scales.

Morphometric Analyses

Patterns of morphological variation were visualized using a semi-strong hybrid multidimensional scaling (SSH MDS) ordination, which was implemented in PATN vers.

3.03 (Belbin 2004; http://www.patn.com.au/). SSH MDS ordination uses a combination of linear and ordinal regression techniques to plot samples in ordination space. A threshold value (default = 0.9) is specified to dictate when linear or ordinal regression methods are used. Several methods were available to evaluate the results of the SSH MDS ordination. First, Analysis of Similarity (ANOSIM) was used to evaluate whether pre-defined groups were significantly different using Monte-Carlo randomization. Second, Principle Component Correlation (PCC) was used to calculate the effect of the input variables and correlation of user defined groups on the ordination, and presents these as vectors. Third, Monte Carlo Attributes in Ordination (MCAO) evaluates the significance of the PCC vectors.

For my analyses, only specimens scored for all four morphometric characters were included (n=195). Sex, larval host plant association and mtDNA group were also included as extrinsic characters. MtDNA group and larval host plant were each examined as grouping variables to see whether these groups correlated to ordinal structuring based on the morphometric data. The user-defined groups were as follows: MtDNA group, Group 1 = unknown, Group 2 = mtDNA Clade 3-1, and Group 3 = mtDNA Clade 3-2; Larval host plant, Group 1 = unknown host, Group 2 = *Picea* sp., Group 3 = *Abies* sp. or *Pseudotsuga menziesii* (Mirbel) Franco, Group 4 = other. Pairwise distances were generated using a Gower metric (Gower 1971), which automatically standardized all measurements to ensure equal weighting. A 2D SSH ordination was performed with 100 random sets. For PCC, all intrinsic and extrinsic variables, as well as the user-defined groups were included. Evaluation with ANOSIM and MCAO were conducted on rows, each with 1000 permutations.

Hybrid Index Scores

A hybrid index score was developed to examine the diagnostic power of the morphometric characters examined in this study. Mean morphometric character measurements were determined for all specimens belonging to nested mtDNA clades 3-1 and 3-2. A two-tailed t-test was used to compare the mean of each morphometric measurement for the two nested mtDNA clades. A midpoint between each pair of means was designated, and values above and below the midpoint were assigned a score of either

+1 or -1. Unknown values are given a value of 0. A single numerical hybrid index score based on the morphometric character scoring system was given to all individuals assigned to mtDNA clade 3-1 and 3-2, and to all museum specimens previously identified as either *D. reniculelloides* and *D. pseudotsugella*. Larval host plant association was then included with the morphometric characters (-1: *Abies* sp. and *P. menziesii*; +1: *Picea* sp.) and a second numerical hybrid index score was assigned to each individual. Accuracy of the hybrid index scores were assessed where a positive value represents clade 3-2 (*D. reniculelloides*), and a negative value represents clade 3-1 (*D. pseudotsugella*).

Results

Sequence Summary Statistics

mtDNA

The 475 bp COI marker was obtained for 96 individuals from 22 localities, representing 22 unique COI haplotypes (Table 4-2). I found 25 variable nucleotide sites (5, 0, 20 changes at first, second and third codon positions respectively), representing 22 synonymous and 3 nonsynonymous changes. Mean base pair frequencies were A: 0.314, C: 0.129, G: 0.162, T: 0.396 and uncorrected pairwise divergence ranged from 0.21% - 2.52%.

Nuclear

Sequence for the nuclear marker ITS2 was obtained for 29 individuals from 13 localities, with at least a single representative from each of the 22 mtDNA haplotypes (Tables 4-2 and 4-3). ITS2 fragment length was variable, either 499 or 491 bp, due to the presence of three indels (Table 4-3). Four unique genotypes were present and were defined by two variable sites (one site polymorphic for A/G) and the three indels (Table 4-3). Sequence for the nuclear marker EF1α was obtained from 33 individuals from 15 localities, representing 18 of 22 mtDNA haplotypes (Tables 4-2 and 4-4). Sequence from the remaining four haplotypes could not be obtained due to degraded DNA. Four apparent homozygous genotypes were present (Table 4-4), defined by three variable sites,

all representing transitions. The absence of heterozygotes was inferred only on the basis of the absence of consistent double peaks in chromatograms for any sequences.

Phylogenetic and Population Genetic Analyses

Parsimony

A heuristic parsimony search on the 22 unique COI haplotypes resulted in 16 most parsimonious trees 96 steps in length (Fig. 4-3). Parsimony analysis was unable to fully resolve relationships among the COI haplotypes, although the *schuetzeella* group was monophyletic, albeit with bootstrap support below 50%.

Nested Clade Analysis

A single haplotype network was produced from the 22 COI haplotypes, with a maximum of 9 mutations between any two haplotypes (Fig. 4-4). The network was composed of nine 1-step clades, 5 two-step clades, and two main three-step clades (Fig. 4-4: clade 3-1 and 3-2). The two main clades roughly correspond to a western clade (3-1) and eastern clade (3-2), although populations in Alaska belonged to the eastern clade, and *D. schuetzeella* from Germany was included in the western clade. Haplotype J was the most common haplotype, formed the hub of the 3-2 clade, had the broadest distribution, and is most likely to be the root of the statistical network (Templeton 1998). Two populations had specimens from each of two main clades (Table 4-2: BC and KS). Uncorrected pairwise divergence was 0.21-1.26% in clade 3-1, and 0.21-0.84% in clade 3-2. Divergence between clades was 1.05-2.52%. Although *D. schuetzeella* was only included for comparison, only 0.84% divergence separated this species from other members of clade 3-1. Based on nested contingency analysis (Posada et al. 2000) only the eastern clade (Clade 3-1) showed significant geographic association (χ^2=52.32, P=0.014). Based on the updated inference key from the GeoDis website (http://darwin.uvigo.es/software/geodis.html) the significant structuring in this clade has resulted from restricted gene flow with isolation by distance (result obtained via the following steps in the GeoDis inference key: 1-2-3-4-NO).

Genetic Differentiation

Tests for population structure on the COI haplotypes indicate a significant genetic differentiation (χ^2=268.59, df=228, P=0.034; N_{ST}=0.68). SAMOVA analyses indicate that the populations were divided into three optimal groups: 1) OR, KS, VU, AL, SL, VS, DC, PH, PR; 2) GE; 3) BC, FM, TB, BR, RL, TC, LB, BB, AK, HY, CH, SE. A total of 80.27% of the variation was explained between groups, and only 0.26% between populations within groups, and the remainder within populations (Table 4-5). The overall F_{ST} was significant (F_{ST}=0.81; P<0.001), suggesting that there is restricted gene flow among these groups. Groups 1 and 3 roughly correspond to the two main clades (3-1 and 3-2, respectively) recovered in the nested clade analysis. The SAMOVA groupings differed from the nested clade analysis in that the European population (GE) formed its own group, and populations BC and KS were not shared between groups.

The spatial distribution of ITS2 genotypes supported the large-scale population structure found by the nested clade analysis of COI haplotypes. All ITS2 AA genotypes were found in individuals belonging to the eastern 3-2 clade, while the vast majority of BB, BC and CC genotypes were contained in the western 3-1 clade. A single specimen in clade 3-2 (haplotype N) contained an ITS genotype (BB) that was otherwise only associated with clade 3-1 (Table 4-3), indicating that the ITS2 genotype data was not completely congruent with the COI haplotype network. Genotype diversity of populations within the western clade were variable, where some populations contained three genotypes (AL) while the majority contained only one (Table 4-2, Fig. 4-4).

The distribution of EF1α genotypes, on the other hand, failed to support the population structuring. The great majority of genotypes were aa, and showed little obvious geographic association (Table 4-2). The other three genotypes were found in two populations (BB and FM) that were separated by approximately 600 km. The apparent absence of heterozygotes was unexpected, but it is plausible that the heterozygotes were not detected due to biased PCR amplification.

Morphometric Analysis

Two 2D SSH ordinations were produced to examine how well variation in the morphometric data is reflected in two user-defined groupings: 1) larval host plant (Fig. 4-5), and 2) mtDNA group (ordination not shown). Both ordinations had low levels of

stress (host plant = 0.1515; mtDNA = 0.1516). In the larval host examination, all morphometric variables, and two extrinsic variables had a significant correlation with the SSH 2D ordination (Length, ADS, ASBA, and Mode: P<0.001; Sex: P=0.009; mtDNA Group: P=0.02) (Fig 4-5). Of the four user-defined larval host plant groups in the SSH ordination (Group 1: unknown host; Group 2: *Picea* sp.; Group 3: *Abies* sp. or *P. menziesii*; Group 4: other), two significant pairwise comparisons were found (Group 1 - Group 2: P<0.001; Group 2 - Group 3: P<0.001). In the mtDNA group examination, the extrinsic variable was significantly correlated to the morphometric ordination (Length, ADS, ASBA, and Mode: P<0.001; mtDNA group: P=0.01). Of the three user-defined groups (Group 1: unknown; Group 2: western clade 3-1; Group 3: eastern clade 3-2) only Group 2 – Group 3 was significant (P<0.001). These results indicate that larval host plant association and mtDNA group both correlate significantly to distinct clusters of morphometric data.

Hybrid Index

Means and standard deviations of the four morphometric measurements were calculated for members of clade 3-1 and 3-2, and the means of each measurement were significantly different (P< 0.05) between the two mtDNA clades (Table 4-6). A midpoint between each mean was identified and used as threshold values for the hybrid index score (Table 4-7). The accuracy of the hybrid index for identification was quite variable (Table 4-8). When using only morphometric measurements, the index correctly identified members of clade 3-1 (*D. pseudotsugella*) 67.6% of the time, and clade 3-2 (*D. reniculelloides*) was correctly identified 64.1% of the time (Table 4-8). For previously identified museum specimens, *D. pseudotsugella* was correctly identified 80.3% of the time, while *D. reniculelloides* was only identified 48.6 % of the time. When larval host was included as an index variable, the percentage of correct identifications increased. For the mtDNA clades, correct identification increased to 70.3% and 68.8% for clade 3-1 (*D. pseudotsugella*) and 3-2 (*D. reniculelloides*), respectively (Table 4-8). For previously identified *D. pseudotsugella* and *D. reniculelloides* specimens, correct identifications increased to 87.9% and 77.1%, respectively (Table 4-8).

Discussion

Taxonomic treatments have delimited two Nearctic species in the *schuetzeella* group; a larger, darker, primarily eastern *D. reniculelloides*, and a smaller, paler, western *D. pseudotsugella* (Munroe 1959, Mutuura and Munroe 1973, Neunzig 2003). Based on the original descriptions, North American members of clades 3-1 and 3-2 corresponded to *D. reniculelloides* and *D. pseudotsugella*, respectively. The two mtDNA groups were resolved using estimates of gene flow and genetic differentiation, supporting the traditional hypothesis of two distinct species. Although restricted gene flow by distance was found to be significant in the nested contingency analysis, members of both clades occurred sympatrically in two populations in southern Alberta and British Columbia. In spite of this overlap, these clades maintain their genomic integrity, both in mtDNA and the ITS2 locus. Maintenance of this genomic integrity strongly suggests that these groups represent two distinct species (Sperling 2003). Their sympatric overlap suggests that isolation by distance is not the only barrier restricting gene flow between these two species. Flight period (Abbot and Withgotti 2004, Forister 2005), host preference (Funk et al. 2002, Nosil et al. 2002, Emelianov et al. 2003, Bethenod et al. 2005, Ohshima and Yoshizawa 2006), or mating behavior, such as pheromone differences (Zhu et al. 1997, Dopman et al. 2005, Sheck et al. 2006), could all contribute to maintaining the integrity of sympatric species. However, all sympatric specimens in this study were collected at the same time and place, making it unlikely that flight period or other allochronic barriers exist between these two species. Detailed examinations of host plant association or pheromone preference have not been conducted, and would be important avenues for future work.

Although *D. reniculelloides* and *D. pseudotsugella* maintain their genomic integrity in spite of a sympatric distribution, one clear discordance was observed. One specimen with an eastern mtDNA haplotype had a western ITS2 genotype (Table 4-3). This specimen had a morphological appearance consistent with the eastern clade, suggesting nuclear rather than mtDNA introgression. Since this specimen was found in the region of sympatry (Fig. 4-1), it is most plausibly due to an introgression event, although a retained ancestral polymorphism cannot be ruled out (Funk and Omland

2003). A single case of incongruence between mtDNA and ITS variants is not enough to abandon support for two distinct species, although the due to the reduced surveys of ITS in the regions of sympatry there may be additional instances of incongruence were not detected. If the incongruence is the result of an introgression event, then it indicates that these species are capable of some interbreeding, although this process appears limited in other sympatric populations. If it is the result of retained polymorphism, then the incomplete coalescence of lineages indicates that these two sister taxa are likely to have diverged only recently. Support for recency of speciation is also supported by the EF1α locus, which did not demonstrate variation congruent with the population structure found in the other two loci, even though some variability does exist.

The populations of *Dioryctria* sampled in this study can be considered to belong to two distinct species, but there was a surprising lack of significant geographic structuring within these species, even though many samples were separated by large distances (Fig 4-1). Very few populations had unique mtDNA haplotypes (Table 4-2: PH and KS), although many were only represented by a single individual. A similar degree of variability was found for sequence variation in *D. abietivorella* (Grote) (Roe et al. 2006) and allozyme variation in *D. disclusa* Heinrich (Richmond 1995). These species showed only moderate genetic differentiation among populations, similar to the eastern and western species observed in this study. This suggests that migration among populations is quite high within *Dioryctria*.

The single European specimen of *D. schuetzeella* was initially included in the analysis to provide comparison with a related but geographically distant congener. However, it exhibited a surprising lack of genetic differentiation from the two Nearctic species. Although it was identified as a separate group in the SAMOVA analysis, only 0.84% divergence separated this species from other members of western *D. pseudotsugella*. This level of divergence is less than that separating *D. pseudotsugella* and *D. reniculelloides* (minimum 1.05%). This calls into question the distinctness of *D. schuetzeella* as a separate species, although it is important to note that only a single specimen was available for examination. Additional material from Europe will need to be examined to assess the distinctness of this species from the two North American species.

Accurate species delimitations require data from many sources such as morphological variation, behavioural traits, molecular loci, and geographic location. Not only is it important to accurately delimit species, but it is equally important to identify characters to diagnose distinct species. In addition to the molecular loci, the four morphometric characters also resolved two distinct clusters. These morphological groups were significantly associated with the eastern and western mtDNA clades, as well as larval host plant (Fig. 4-5). The morphometric measurements were based on the original species descriptions of *D. reniculelloides* and *D. pseudotsugella*, supporting the traditional delimitation of these two described species. Unfortunately, due to intraspecific variability of the morphometric characters, they must be used in combination and cannot accurately separate the two species independently (Table 4-8). Even though the means of each morphometric measurement were significantly different between the two mtDNA clades, there was a large overlap in the range of the measurements (Table 4-6). This overlap was also seen in the ordination (Fig. 4-5), and although significantly different, species could not be identified by location in ordination space alone. The hybrid index only achieved moderate success as a method of assignment when morphometric characters alone were used (Table 4-8). Larval host plant associations were considered an important diagnostic character for *D. reniculelloides* and *D. pseudotsugella* in the original descriptions (Munroe 1959, Mutuura and Munroe 1973, Neunzig 2003). When larval host plant was included as a variable in the hybrid index, the identification accuracy increased by as much as 30% (Table 4-8).

In conclusion, molecular (mtDNA and ITS2), morphometric, and larval host plant association were all useful in delimiting *D. reniculelloides* and *D. pseudotsugella*. I found that mtDNA most consistently diagnosed *D. reniculelloides* and *D. pseudotsugella*, supporting general claims of effectiveness by DNA barcoding advocates. However, given the low levels of divergence separating these species, it would be unwise to rely solely on a single character. The use of multiple lines of evidence to delimit and diagnose species has recently been referred to as 'integrative taxonomy' (Dayrat 2005, Will et al. 2005). Using a range of characters to describe species boundaries and identify species provides a much more robust characterization of the species than any single

character system (Rubinoff and Holland 2005), and will provide a deeper understanding of the organism as a whole.

Table 4- 1: Localities for specimens surveyed for molecular and morphometric characters. Latitude (Lat) and longitude (Long) are given in decimal degrees. Number of specimens (N) and rearing information (Host) are also shown.

Locality Code	Locality	Lat	Long	(N)	Host
SL	CAN: BC: 30 km W Summerland	49.720	-120.178	11	unknown
KS	CAN: BC: Vernon: Kalamalka Prov Prk	50.208	-119.269	3	unknown
VS	CAN: BC: Vernon: Silverstar Prov Park	50.344	-119.115	1	unknown
PR	CAN: BC: 10 km S Pritchard	50.599	-119.892	2	unknown
AL	CAN: BC: Adam Lake Prov. Park	50.945	-119.269	6	unknown
BR	CAN: AB: Belly River Cmpgr.	49.025	-113.679	1	*Picea glaucus*
CH	CAN: AB: Cypress Hills	49.665	-110.260	1	*Picea glaucus*
BC	CAN: AB: Beavercreek Prov Rec Area	49.804	-113.935	14	unknown
PH	CAN: AB: Porcupine Hills, 15 km E of Jct of Forestry Trunk Rd and Hwy 514	49.845	-114.264	1	*Pseudotsuga menzeisii*
DC	CAN: AB: Dutch Creek Cmpgr.	49.909	-114.391	1	unknown
VU	CAN: AB: 15 km SE Vulcan	50.358	-113.096	6	unknown
TB	CAN: AB: Tolman Bridge	51.834	-113.008	1	unknown
BB	CAN: AB: 7 mi N Bearberry	51.839	-114.710	8	unknown
TC	CAN: AB: Thompson Cr. Rec Area	52.012	-116.628	4	*Picea glaucus*
SE	CAN: AB: Seibert Lk	54.692	-111.265	2	unknown
HY	CAN: AB: Hythe	55.331	-119.455	1	*Picea glaucus*
FM	CAN: AB; Fort MacMurray	56.736	-111.378	16	unknown
RL	CAN: AB: Rainbow Lk	58.297	-119.404	1	*Picea glaucus*
LB	CAN: AB: La Butte Prov Prk	59.376	-111.201	2	unknown
OR	USA: OR: Benton Co.: Corvallis	44.564	-123.262	4	unknown
AK	USA: AK: Anchorage Co.: Anchorage	61.218	-149.900	8	*Picea pungens*
GE	Germany: Saxony	49.000	12.1000	1	unknown

Table 4- 2: Sample locality (for abbreviations, see Table 4-1), sample size (N), haplotype and genotype information for specimens surveyed in this study. Numbers in brackets indicate number of individuals for haplotypes or genotypes, if there are multiple specimens per locality.

Locality	COI				ITS2		EF1α	
	N	Haplotype	π^a	h^b	N	Genotype	N	Genotype
SL	11	A(9) H I	0.0011	0.35	1	BC	1	aa
KS	3	A(2) N	0.0114	0.67	1	BB	1	aa
VS	1	A	-	-	-	-	1	aa
PR	3	A(2) E	0.0014	0.67	-	-	-	-
AL	6	A(3) B D G	0.0021	0.80	4	BB, BC CC(2)	2	aa
BR	1	J	-	-	-	-	-	-
CH	1	Q	-	-	-	-	-	-
BC	14	A(2) J(5) L Q(4) T(2)	0.0055	0.80	4	AA(3) BB	4	aa
PH	1	F	-	-	1	CC	1	aa
DC	1	H	-	-	1	BC	1	aa
VU	6	A(2) E G(2) H	0.0032	0.87	2	BB, BC	2	aa
TB	1	J	-	-	-	-	1	aa
BB	8	J(2) M P R Q(2) U	0.0038	0.93	5	AA	5	aa(3) bb, cc,
TC	4	J(2) Q(2)	0.0014	0.67	-	-	-	-
SE	2	J	0.0000	0.0	-	-	-	-
HY	1	Q	-	-	1	AA	1	aa
FM	16	J(10) K O Q(4)	0.0015	0.58	3	AA	6	aa(4) bb, dd
RL	1	J	-	-	-	-	-	-
LB	2	J Q	0.0021	1.0	1	AA	1	aa
OR	4	A(2) C H	0.0032	0.83	1	BC	2	aa
AK	8	J(3) O Q(3) S	0.0226	0.79	4	AA	4	aa
GE	1	V	-	-	-	-	-	-
Total (Average)	96		0.0087	0.82	29		33	

[a] nucleotide diversity (Nei 1987 equation 10.5)
[b] haplotype diversity (Nei 1987 equation 8.4)

Table 4- 3: Nucleotide variation and genotype designation for 29 individuals sequenced for ITS2. For locality (LC) abbreviations, see Table 4-1. Invariant positions are indicated with an apostrophe (‘), and gaps in the alignment are indicated with a dash (-). For reference, COI haplotypes (COI Hap) and the location of each specimen in the two main mtDNA clades is indicated (NCA). The specimen in bold highlights the presence of a discordant combination of haplotypes and genotypes.

Specimen Number	LC	134	135	156	247	253	254	255	256	257	403	ITS Gen	COI Hap	NCA
AR72	LB	G	C	T	T	C	A	G	G	C	G	AA	J	3-2
AR73	FM	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	J	3-2
AR75	FM	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	J	3-2
AR77	FM	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	K	3-2
AR274	BC	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	J	3-2
AR309	BB	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	U	3-2
AR311	BB	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	M	3-2
AR359	HY	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	Q	3-2
D01	BB	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	P	3-2
D04	BB	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	J	3-2
D07	BB	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	R	3-2
D35	BC	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	T	3-2
D36	BC	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	L	3-2
D57	AK	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	O	3-2
D58	AK	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	S	3-2
D59	AK	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	Q	3-2
D61	AK	‘	‘	‘	‘	‘	‘	‘	‘	‘	‘	AA	J	3-2
AR105	OR	-	-	-	C	-	-	-	-	-	A/G	BC	C	3-1
AR208	VU	-	-	-	C	-	-	-	-	-	G	BB	E	3-1
AR210	DC	-	-	-	C	-	-	-	-	-	A/G	BC	H	3-1
AR276	BC	-	-	-	C	-	-	-	-	-	G	BB	A	3-1
AR314	AL	-	-	-	C	-	-	-	-	-	A	CC	G	3-1
AR360	PH	-	-	-	C	-	-	-	-	-	A	CC	F	3-1
AR366	**KS**	**-**	**-**	**-**	**C**	**-**	**-**	**-**	**-**	**-**	**G**	**BB**	**N**	**3-2**
D09	VU	-	-	-	C	-	-	-	-	-	A/G	BC	A	3-1
D26	AL	-	-	-	C	-	-	-	-	-	A	CC	B	3-1
D28	AL	-	-	-	C	-	-	-	-	-	A/G	BC	A	3-1
D29	AL	-	-	-	C	-	-	-	-	-	G	BB	D	3-1
D45	SL	-	-	-	C	-	-	-	-	-	A/G	BC	I	3-1

Table 4- 4: Nucleotide variation and genotype designation for 33 individuals sequenced for EF1α. For locality (LC) abbreviations, see Table 4-1. Nucleotide position is shown relative to *D. melanogaster* (Hovemann et al 1988). Invariant positions are indicated with an apostrophe (').

Specimen Number	LC	2431	2443	2503	EF1α Genotype
Majority of Specimens (n=29)		C	C	A	aa
D03	BB	T	'	'	bb
D24	FM	T	'	'	bb
AR73	FM	'	T	G	dd
AR309	BB	'	T	'	cc

Table 4- 5: SAMOVA results for genetic differentiation between three optimal groups roughly corresponding to mtDNA network clades 2-1+2-2, 2-3, and 3-2.

Source of variation	Variance components	Percent variation
Among groups	2.86	80.27
Among populations within groups	0.01	0.26
Within populations	0.69	19.47
Overall F_{ST}	0.81*	

* Significant at 0.05 level.

Table 4- 6: Distribution of four morphometric measurements in the two main mtDNA clades found in the nested clade analysis. All measurements were taken in ImageJ. A comparison of the morphometric character means was conducted using a two-tailed t-test.

Morphometric Character		Nested Clade 3-1	Nested Clade 3-2	t-test
Length (mm)	Range:	9.2-11.8	9.9-11.9	
	Mean (± st. dev.)	10.2 (±0.7)	10.9 (±0.5)	*
ADS (mm^2)[a]		16.8-27.7	20.6-34.3	
		21.6 (±3.1)	26.4 (±3.7)	*
ASBA (mm^2)[b]		0.3-2.2	0.1-1.3	
		1.1 (±0.4)	0.7 (±0.3)	*
SBA Mode[c]		62-138	54-118	
		109.5 (±17.8)	90.3 (±19.4)	*

[a]Area of dark scales in forewing
[b]Area of the pinkish-orange patch in the subbasal area of the forewing
[c]Modal scale colour (in grayscale) of the patch in the subbasal area
*$P<0.05$

Table 4- 7: Morphometric and larval host characters used for scoring a hybrid index for individuals identified as *D. reniculelloides* and *D. pseudotsugella* in mtDNA clades or previous taxonomic identifications. Negative values were assigned to characters that were *D. pseudotsugella*-like, and conversely positive for *D. reniculelloides*-like characters based on the original descriptions and diagnoses of the two species.

Character	State	Score
Host	*Abies* sp. or *Pseudotsuga menziesii*	-1
	Picea sp.	1
	other or unknown	0
Forewing Length	≤ 10.5 mm	-1
	> 10.5 mm	1
	unknown	0
Area of Dark Scales	≤ 24.0 mm^2	-1
	> 24.0 mm^2	1
	unknown	0
Area of Pale Subbasal Patch	≥ 0.9 mm^2	-1
	> 0.9 mm^2	1
	unknown	0
Modal Greyscale Colour of Subbasal Area	≥ 99.9	-1
	> 99.9	1
	unknown	0

Table 4- 8: Frequency of correct and incorrect identifications using two hybrid indices: 1) morphometric characters only, or 2) morphometric characters and larval host association. Indices were evaluated against specimens classified on the basis of mtDNA clade, and museum specimens identified prior to this investigation. A negative score indicates *D. pseudotsugella*-like characters, while a positive score indicates *D. reniculelloides*-like characters (for character descriptions see Table 4-7).

	Including Host		Excluding Host	
	+ Score	- Score	+ Score	- Score
MtDNA clade 3-1	16.2%	70.3%	16.2%	67.6%
MtDNA clade 3-2	68.8%	15.6%	64.1%	18.8%
D. pseudotsugella	7.6%	87.9%	7.6%	80.3%
D. reniculelloides	77.1%	15.6%	48.6%	21.2%

Figure 4- 1: Locations of specimens examined. Filled circles indicate specimens used for both molecular (DNA) and morphological (morph) surveys, and open circles indicate localities of specimens surveyed for morphological characters alone. Three locations not shown are sites in Anchorage AK (DNA+morph), Timmins ON (morph), and Germany (DNA+morph). The shaded area indicates the region of sympatric distribution, with dashed line separating the putative ranges of *Dioryctria reniculelloides* (northeastern distribution) and *D. pseudotsugella* (southwestern distribution).

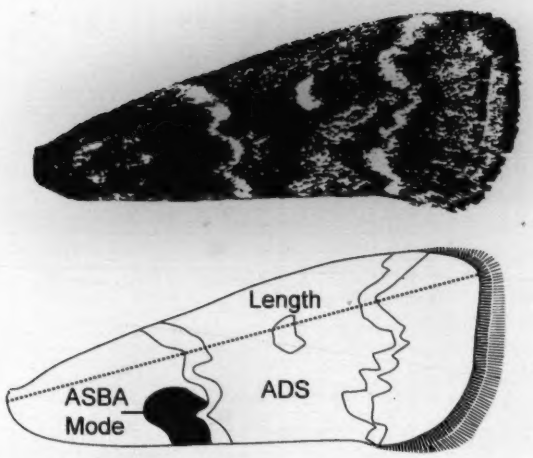

Figure 4-2: Forewing measurements used for morphometric analysis of morphological variation. Length: Maximum forewing length, not including fringe. ASBA: Area of the pinkish-orange subbasal patch. Mode: Median scale color of pinkish-orange subbasal patch. ADS: Area of dark scales for the entire surface of the forewing, excluding fringe.

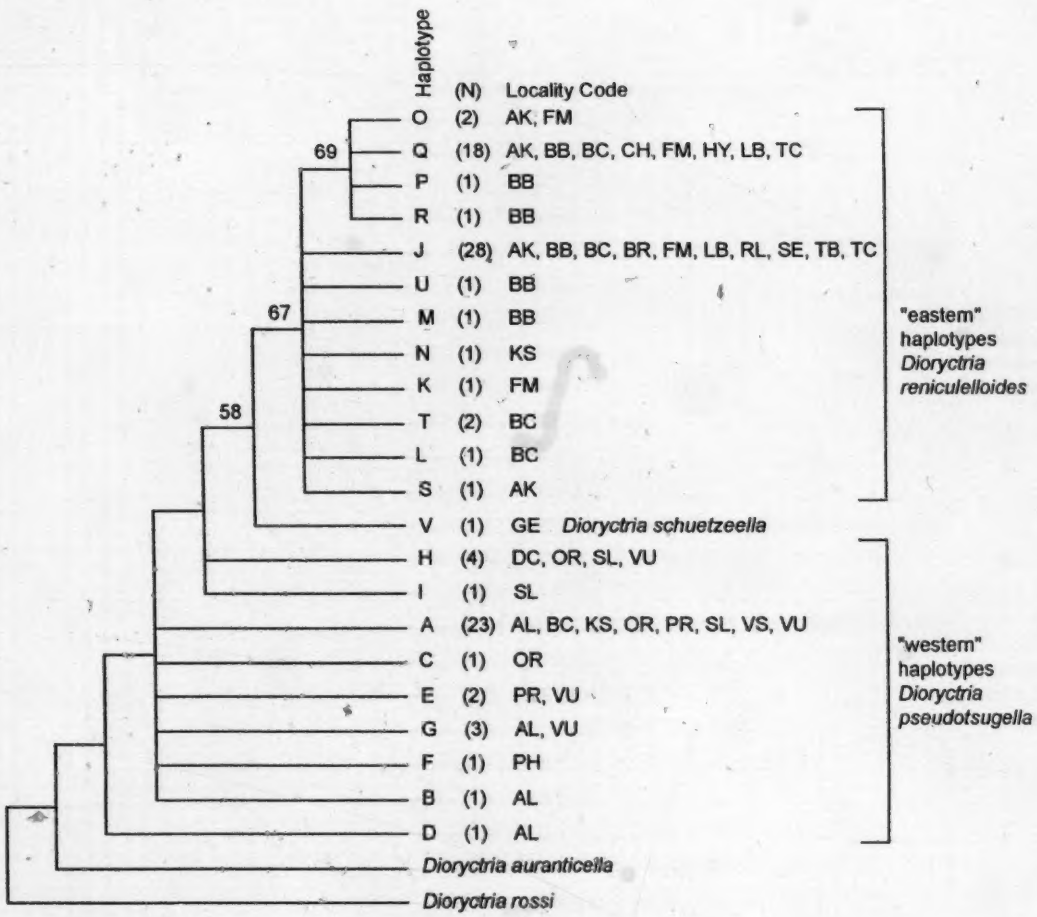

Figure 4- 3: Strict consensus tree of 16 most parsimonious trees (length = 96) of 22 COI haplotypes representing 96 specimens from 22 localities in Alberta, British Columbia, Alaska and Oregon. Branch support is indicated by bootstrap values if over 50%. COI haplotype, number of specimens and collecting locality are indicated on terminal nodes. *D. auranticella* and *D. rossi* were used as outgroup taxa.

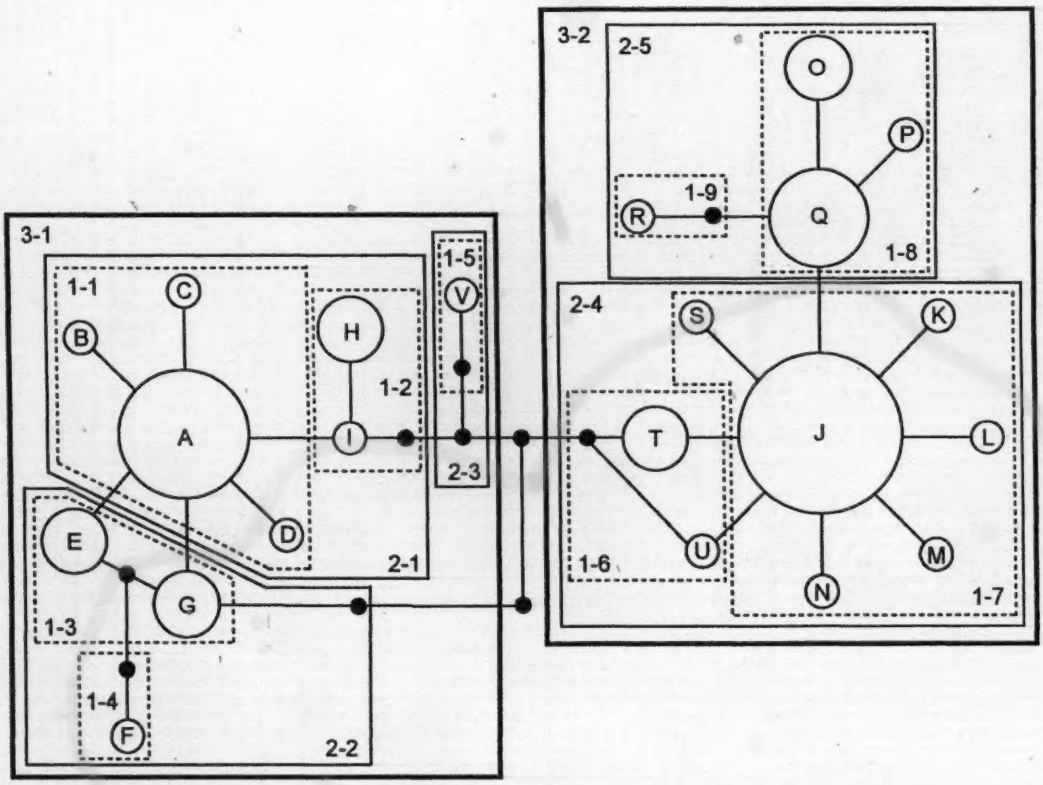

Figure 4- 4: A nested parsimony network of 22 COI haplotypes representing 96 individuals. Hollow circles represent individual haplotypes, with size proportional to the number of specimens. Solid circles indicate missing haplotypes, based on the parsimony analysis. Lines linking haplotypes represent single nucleotide changes. Dashed numbers indicate the nesting pattern of the haplotypes, with dashed boxes representing one-step clades, thin solid lines are two-step clades, and dark solid lines are three-step clades. Haplotype J was found to have the greatest root possibility, based on parsimony analysis.

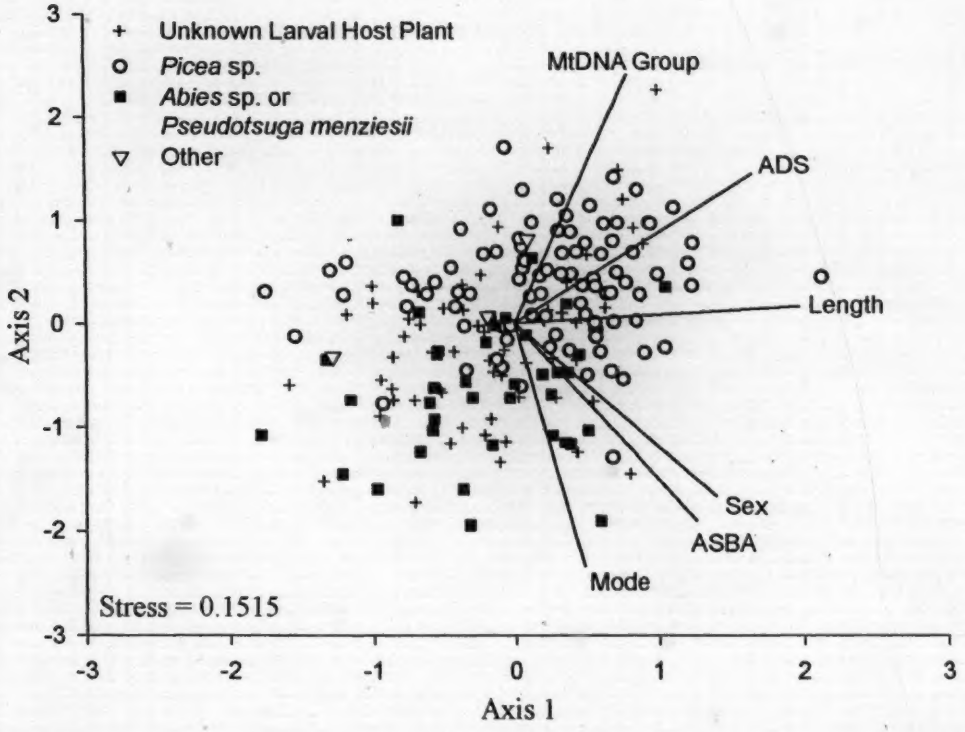

Figure 4- 5: A semi-strong hybrid multidimensional ordination (SSH MDS) of four morphometric forewing characters: Length (maximum forewing length), ADS (area dark scales), ASBA (area of the pinkish-orange subbasal area), and Mode (modal grayscale colour value of scales in the subbasal area). Three extrinsic characters were also overlain as vectors (Sex, mtDNA group, and host plant). The ordination showed significant clustering when grouped by host plant and all vectors shown are also significant at P<0.05.

Literature Cited

Abbot, P., and J. H. Withgotti. 2004. Phylogenetic and molecular evidence for allochronic speciation in gall-forming aphids (*Pemphigus)*. Evolution 58: 539-553.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18: 489-522.

Belbin, L. 2004. PATN computer program, version 3.03. By Belbin, L., Canberra, Australia.

Bethenod, M.-T., Y. Thomas, F. Rousset, B. Frerot, L. Pelozuelo, G. Genestier, and D. Bourguet. 2005. Genetic isolation between two sympatric host plant races of the European corn borer, *Ostrinia nubilalis* Hubner. II: assortative mating and host-plant preferences for oviposition. Heredity 94: 264-270.

Clary, D. O., and D. R. Wolstenholme. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. Journal of Molecular Evolution 22: 252-271.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657-1659.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Dayrat, B. 2005. Towards integrative taxonomy. Biological Journal of the Linnean Society 85: 407-415.

Dopman, E., L. Perez, S. M. Bogdanowicz, and R. G. Harrison. 2005. Consequences of reproductive barriers for genealogical discordance in the European corn borer. Proceedings of the National Academy of Sciences of the United States of America 102: 14706-14711.

Dupanloup, I., S. Schneider, and L. Excoffier. 2002. A simulated annealing approach to define the genetic structure of populations. Molecular Ecology 11: 2571-2581.

Emelianov, I., F. Marec, and J. Mallet. 2003. Genomic evidence for divergence with gene flow in host races of the larch budmoth. Proceedings of the Royal Society of London Series B 652: 1-9.

Forister, M. L. 2005. Influence of host plant phenology on *Mitoura nelsoni* (Lepidpotera: Lycaenidae). Annals of the Entomological Society of America 98: 295-301.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution and Systematics 34: 397-423.

Funk, D. J., K. E. Filchak, and J. L. Feder. 2002. Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116: 251-267.

Godfray, H. C. J. 2002. Challenges for taxonomy. Nature 417: 17-19.

Gower, J. 1971. A general coefficient of similarity and some of its properties. Biometrics 27: 857-871.

Hebert, P. D. N., and T. R. Gregory. 2005. The promise of DNA barcoding for taxonomy. Systematic Biology 54: 852-859.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological Identification through DNA barcodes. Proceedings of the Royal Society of London Series B 270: 313-321.

Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemiak, and C. M. Francis. 2004. Identification of birds through DNA barcodes. Public Library of Science Biology 2: e312.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2005. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the United States of America 101: 14812-14817.

Hovemann, B., S. Richter, U. Walldorf, and C. C. 1988. Two genes encode related cytoplasmic elongation factor 1 alpha (EF-1a) in *Drosophila melanogaster* with continuous stage specific expression. Nucleic Acids Research 16: 3175-3194.

Hudson, R. R., D. D. Boos, and N. L. Kaplan. 1992. A statistical test for detecting geographic subdivision. Molecular Biology and Evolution 9: 138-151.

Knolke, S., S. Erlacher, A. Hausmann, M. A. Miller, and A. H. Segerer. 2005. A procedure for combined genitalia dissection and DNA extraction in Lepidoptera. Insect Systematics & Evolution 35: 401-409.

Laffin, R. D., D. W. Langor, and F. A. H. Sperling. 2004. Populations structure and gene flow in the white pine weevil, *Pissodes strobi* (Coleoptera: Curculionidae). Annals of the Entomological Society of America 97: 949-956.

Lynch, M., and T. J. Crease. 1990. The analysis of population survey data on DNA-sequence variation. Molecular Biology and Evolution 7: 377-394.

Mallet, J., and K. Willmot. 2003. Taxonomy: Renaissance or Tower of Babel? Trends in Ecology and Evolution 18: 57-59.

Meyer, C. P., and G. Paulay. 2005. DNA barcoding: error rates based on comprehensive sampling. Public Library of Science, Biology 3: e422.

Munroe, E. 1959. Canadian species of *Dioryctria* Zeller (Lepidoptera: Pyralidae). The Canadian Entomologist 91: 65-72.

Mutuura, A., and E. Munroe. 1973. American species of *Dioryctria* (Lepidoptera: Pyralidae) IV. The *schuetzeella* group and the taxonomic status of the spruce cone moth. The Canadian Entomologist 105: 653-668.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Neunzig, H. H. 2003. Pyraloidea, Pyralidae (part), Phycitinae (part). The Moths of America North of Mexico fasc. 15.5.

Nosil, P., B. J. Crespi, and C. P. Sandoval. 2002. Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature 417: 440-443.

Ohshima, I., and K. Yoshizawa. 2006. Multiple host shifts between distantly related plants, Juglandaceae and Ericaceae, in the leaf-mining moth *Acrocercops leucophaea* complex (Lepidoptera: Gracillariidae). Molecular Phylogenetics and Evolution 38: 231-240.

Posada, D., K. A. Crandall, and A. R. Templeton. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Molecular Ecology 9: 487-488.

Rasband, W. 2006. ImageJ computer program, version 1.36b. By Rasband, W., Bethesda, Maryland USA.

Reed, R. D., and F. A. H. Sperling. 1999. Interaction of process partitions in phylogenetic analysis: An example from the swallowtail butterfly genus *Papilio*. Molecular Biology and Evolution 16: 286-297.

Richmond, J. A. 1995. Macrogeographic genetic variation in populations of the webbing coneworm (Lepidoptera: Pyralidae). Journal of Entomological Science 30: 349-358.

Roe, A. D., J. D. Stein, N. E. Gillette, and F. A. H. Sperling. 2006. Identification of *Dioryctria* (Lepidoptera: Pyralidae) in a seed orchard at Chico, California. Annals of the Entomological Society of America 99(3) 433-448.

Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.

Rubinoff, D., and B. S. Holland. 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Systematic Biology 54: 952-961.

Sheck, A., A. Groot, C. Ward, C. Gemeno, J. Wang, C. Brownie, C. Schal, and F. Gould. 2006. Genetics of sex pheromone blend differences between *Heliothis virescens* and *Heliothis subflexa*: a chromosome mapping approach. Journal of Evolutionary Biology 19: 600-617.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87: 651-701.

Sites, J. W., and J. C. Marshall. 2003. Delimiting species: a Renaissance issue in systematic biology. Trends in Ecology & Evolution 18: 462-470.

Sopow, S. L., R. G. Bennet, J.-F. Landry, and B. Landry. 1996. Identification of the "grey' *Dioryctria* species of British Columbia (Lepidoptera: Pyralidae). Journal of the Entomological Society of British Columbia 93: 75-91.

Sperling, F. 2003. Butterfly molecular systematics: from species definitions to higher-level phylogenies, pp. 431-458. *In* C. Boggs, W. Watt and P. Ehrlich [eds.],

Butterflies: Evolution and Ecology Taking Flight. University of Chicago, Chicago Press, Chicago.

Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). computer program, version 4. By Swofford, D. L., Sunderland, Massachusetts.

Templeton, A. R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7: 381-397.

Templeton, A. R. 2001. Using phylogeographic analyses of gene trees to test species status and processes. Molecular Ecology 10: 779-791.

Templeton, a. R., and C. F. Sing. 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. nested analyses with cladogram uncertainty and recombination. Genetics 134: 659-669.

Templeton, a. R., E. Routman, and C. a. Phillips. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. Genetics 140: 767-782.

White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA gene for phylogenetics, pp. 482. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White [eds.], PCR Protocols: A Guide to Methods and Applications. Academic Press Inc., San Diego CA.

Will, K. W., and D. Rubinoff. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20: 47-55.

Will, K. W., B. D. Mishler, and Q. D. Wheeler. 2005. The perils of DNA barcoding and need for integrative taxonomy. Systematic Biology 54: 844-851.

Wilson, E. O. 2003. The encyclopedia of life. Trends in Ecology and Evolution 18: 77-80.

Zhu, J., B. Chastain, B. Spohn, and K. Haynes. 1997. Assortative mating in two pheromone strains of the cabbage looper moth, *Trichoplusia ni.* Journal of Insect Behavior 10: 805-817.

Chapter 5: General Conclusions

Designation and application of species limits is necessary for many fields of biology (Greene 1997, Meyerson and Reaser 2002, Besansky et al. 2003, Armstrong and Ball 2005, Balakrishnan 2005), so robust, well-supported species delimitations are essential. Throughout this thesis I explored a number of aspects to species delimitation. In Chapter 2, I used mtDNA, morphological, pheromone, and larval host variation to identify eight distinct *Dioryctria* species in a seed orchard in Chico CA, and developed a series of diagnostic field characters to identify species in this area. In Chapter 3, I showed that mtDNA nucleotide substitution patterns are highly heterogeneous within and between species, and that longer sequence fragments are required to ensure consistent divergence estimates when relying on a single mtDNA fragment to identify and delimit species. In Chapter 4, I resolved two concordant mtDNA and nuclear clades representing two distinct *Dioryctria* species. These species are also distinguished by morphological and behavioural differences, despite overlapping morphological variation and geographic ranges.

Inferring species boundaries is not always straightforward. Boundaries between species are often fuzzy and lack distinct concordance between character sets (Sites and Marshall 2004, Cardoso and Vogler 2005). Ambiguous boundaries can occur between species that are sympatric, morphologically cryptic, hybridizing or undergoing rapid diversification (Salzburger et al. 2002, Shaw 2002, Bettles et al. 2005, Cardoso and Vogler 2005, Marko 2005, Barluenga et al. 2006). Species are often assumed to be monophyletic (Harrison 1998), but theoretical models of speciation indicate that monophyly is not instantaneous, rather species progress through stages of polyphyly and paraphyly before monophyly is reached (Avise and Ball 1990, Avise 2000), making species delimitations challenging (Funk and Omland 2003). Here I will discuss several frequently encountered situations that can cause difficulties and errors when delimiting species.

Sampling Bias

Insufficient sampling is often cited as a major reason for inaccurate species delimitations (Funk 1999, Morando et al. 2003, Wahlberg et al. 2003, Sites and Marshall 2004). Limited specimen sampling reduces the probability of detecting incomplete lineage sorting, cryptic diversity, or hybridization events, particularly when these phenomena occur in sympatry (see below). This may lead to an oversimplified view of species limits and provide misleading, albeit well supported species boundaries by failing to recover complex genetic variation caused by localized divergence, population bottlenecks, random genetic drift, or hybridization (Funk 1999, Wiens and Penkrot 2002).

Like specimens, characters (i.e. morphological, molecular, or ecological) are also sensitive to biased sampling (Morando et al. 2003, Sites and Marshall 2004, Hendrixson and Bond 2005, Meyer and Paulay 2005). Many studies rely on a single marker to infer species boundaries (Hebert et al. 2003, Hebert et al. 2005), which can lead to under- or overestimations of species diversity, or fail to detect incongruence (Paquin and Hedin 2004, Will and Rubinoff 2004, Meyer and Paulay 2005, Strand and Sundberg 2005).

Incongruence

As mentioned above, incongruence is another major pitfall for species delimitation, and occurs when independent data sets resolve conflicting species limits. A variety of events can lead to incongruence (Avise 2000), and I will expand and discuss three of the most common causes: 1) incomplete lineage sorting, 2) introgressive hybridization, and 3) nuclear pseudogenes.

Incomplete Lineage Sorting

Incomplete lineage sorting, or retained ancestral polymorphism, occurs when alleles have not reached fixation within their respective lineages and are shared between two separate lineages (Nichols 2001, Funk and Omland 2003). Coalescence times can vary on several different levels, such as between lineages, among genes or within genes. Variation in coalescent times for different loci can result in incongruent gene trees, and is

particularly problematic when species are recently diverged, or when the loci have large effective populations (Pamilo and Nei 1988, Takahata 1989, Maddison 1997, Edwards and Beerli 2000). As divergence between species increases, more ancestral polymorphisms are lost to random stochastic processes, until all gene trees are isomorphic with the actual species tree (Avise 2000, Sites and Marshall 2004). MtDNA has a much smaller population size than most nuclear loci, so coalescent times for mtDNA genes are less than most nuclear loci (Moore 1995, Nichols 2001, Hendrixson and Bond 2005). Moreover, mtDNA may reach fixation prior to the rest of the genome and provide a misleading, simplistic view of the species tree and boundaries (Sites and Marshall 2004).

Coalescent times are usually discussed with respect to molecular characters, but morphological, ecological and behavioural characters can also coalesce at variable rates. Many morphological characters, such as forewing pattern (Chapter 2: *D. pentictonella*) or size (Chapter 4) can be highly variable and prone to regional or environmental pressures (Price et al. 2003, Braendle et al. 2005, Brisson et al. 2005, Gubitz et al. 2005, Marko 2005). Ecological or behavioural characters, such as pheromone or host races, can also be variable, with race formation often occurring prior to a speciation event (Emelianov et al. 2003, Diegisser et al. 2004, Bethenod et al. 2005, Blair et al. 2005). Some structures, such as male genitalia, are considered very good sources of characters for delimiting species in some groups, but reliability of characters can vary among lineages, often dependent on the level of sexual selection (Eberhard 1985, Arnqvist 1997, Mutanen 2005).

Introgression and interspecific hybridization

Introgression is the movement of foreign genes into a genome, and is most commonly caused by interspecific hybridization. Interspecific hybridization acts as a sieve for gene flow, providing a means of introduction of foreign gene material into a genome (Funk and Omland 2003, Mallet 2005). Introgressed loci will have inaccurate gene trees and provide well supported species boundaries and relationships, but will be incongruent with the actual species tree (Shaw 2002, Wahlberg et al. 2003, Bensch et al. 2006, Sanders et al. 2006). Although hybridization and introgression have been

overlooked and discounted in the past (Mallet 2005), these processes give me insight into the history of an organism and a better understanding of the evolutionary forces shaping a species. Moreover, interspecific hybridization has been shown to cause speciation in a range of animal taxa (Dowling and Secor 1997, Donnelly et al. 2004, Mallet 2005), making it an important contributor to overall species diversity.

Nuclear Pseudogenes

Nuclear encoded mitochondrial pseudogenes (numts) are non-coding copies of mitochondrial genes that are often mistaken for actual mtDNA sequence (Bensasson et al. 2001). If numts are used in place of the true mtDNA sequence, these paralogous genes could provide completely misleading and inaccurate gene trees (Zhang and Hewitt 1996, Moritz and Cicero 2004, Thalmann et al. 2005).

Recommendations

Throughout this thesis, I have demonstrated that a wide range of factors can contribute to inaccurate species limits and difficulties when defining species boundaries. To minimize the effects of these factors, several recommendations should be followed. First and foremost, broad sampling of populations, species and geographic range must be conducted. Realistically, complete sampling can never be obtained, but it should be representative of the biological variation existing within and between species (Funk 1999, Wahlberg et al. 2003, Olsson et al. 2005). Multiple lines of evidence should be used to validate and cross-validate species limits (Sites and Marshall 2004). Large data sets are needed, particularly when trying to span intra- and interspecific levels of divergence. For example, in Chapter 2, I sampled specimens from a range of host plants, flight times, morphological variation, and pheromone types to ensure that all possible biological variation was examined. If I had targeted only pheromone and host plant variation, several sympatric species would not have been uncovered. Similarly in Chapter 4, the sympatric distribution of *D. reniculelloides* and *D. pseudotsugella* would not have been recorded if localities were not properly sampled. The importance of intra and interspecific sampling was well demonstrated in Chapter 3; mtDNA divergence overlap

among *Anopheles* species would not have been observed and could have led to inaccurate estimations of species limits if intra- and interspecific divergences were not examined.

Although some problems such as insufficient sampling can be rectified, it is important to note that hybridization, variable rates of evolution and retained polymorphisms are real biological events, not necessarily 'bad' speciation or taxonomy. Rather, these events provide insight into the evolutionary processes occurring between closely related species or between populations in the process of diverging. Incongruence can indicate hidden, morphologically cryptic taxa, or suggest modes of speciation, and when seen in the right light, is often more interesting than complete congruence (Wiens and Penkrot 2002). To ensure that these processes are documented, multiple data sources again are needed (Pamilo and Nei 1988, Moore 1995, Nichols 2001, Wahlberg et al. 2003, Balakrishnan 2005). Incongruence between gene trees and species trees can only be detected when multiple loci are used, especially when one locus is mtDNA (Shaw 2002, Ballard and Whitlock 2004, Sanders et al. 2006) as was demonstrated in Chapter 4 when I was able to identify a single introgression/incomplete lineage sorting event in *D. reniculelloides*. Of course, inclusion of ecological and morphological data is also important. As appropriately stated by Kiefer et al (2002, pg. 557): "molecular phylogenetic analyses do not free systematists from a thorough inclusion of morphological and ecological data". This is especially true given the support for projects such as DNA barcoding which relies solely on a single molecular marker to infer species boundaries. Morphological and ecological data, when used in conjunction with molecular data, are more powerful tools for delimiting species and provide a better insight into the organism, than molecular data alone (Moore 1995, Lipscomb et al. 2003, Mallet and Willmot 2003, Dayrat 2005).

Future Directions

The genus *Dioryctria* has proven to be a group with enormous research potential. This genus is highly diverse, and found throughout the world, although the bulk of the diversity occurs in North America. Most of the Nearctic diversity is found in the "raised scale" species groups which feed primarily on the genus *Pinus*. Many species are

difficult to diagnose and species show little molecular divergence. I feel that these groups need to be examined in greater detail, with a similar approach to that conducted with *D. reniculelloides* and *D. pseudotsugella.* Even more interesting is the restricted host plant use characterized by each species. An in-depth examination of larval host plant associations among members of the raised-scale species groups would help to understand why these groups are so diverse.

The relationships among the species groups of *Dioryctria* are also very interesting. Du et al. (2005) published a study demonstrating some of these relationships, but many basal relationships were poorly supported, and some species groups were poorly sampled or missing completely. Broader taxon sampling and use of more conserved molecular loci such as EF1α are needed to expand on our current understanding of the phylogenetic relationships among *Dioryctria* (Du et al. 2005).

Specimens of *D. pentictonella* collected by pheromone trapping showed an impressive range of forewing variability (Chapter 2). This variability did not appear linked to flight period, pheromone or host. Further exploration of the source of forewing variability and the factors affecting it could help to provide a better understanding of morphological variability (or lack thereof) in *Dioryctria*, and help to identify reliable diagnostic characters.

Over the course of this thesis, I had the opportunity to explore the molecular evolution of species differences. Several areas within this field warrant further study. First, more rigorous testing of mutation hotspots within COI-COII would prove highly informative. Understanding how and where mutations accumulate, particularly in third base pair positions is very important for understanding the development of species differences. Second, this line of investigation should be expanded to examine differences in location between synonymous and nonsynonymous changes, particularly with respect to saturation and loss of phylogenetic signal. Expanding our understanding of the processes of mtDNA evolution is especially important in light of the growing popularity of DNA taxonomy and DNA barcoding to identify and delimit species.

Literature Cited

Armstrong, K. F., and S. L. Ball. 2005. DNA barcodes for biosecurity: invasive species identification. Philisophical Transactions of the Royal Society of London Series B 360: 1813-1823.

Arnqvist, G. 1997. The evolution of animal genitalia: distinguishing between hypotheses by single species studies. Biological Journal of the Linnean Society 60: 365-379.

Avise, J. C. 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

Avise, J. C., and R. M. Ball. 1990. Principles of Genealogical Concordance in Species Concepts and Biological Taxonomy. Oxford Surveys in Evolutionary Biology: 45-67.

Balakrishnan, R. 2005. Species concepts, species boundaries and species identification: a view from the tropics. Systematic Biology 54: 689-693.

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729-744.

Barluenga, M., K. N. Stolting, W. Salzburger, M. Muschick, and A. Meyer. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439: 719-723.

Bensasson, D., D. X. Zhang, D. L. Hartl, and G. M. Hewitt. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecology & Evolution 16: 314-321.

Bensch, S., D. E. Irwin, J. H. Irwin, L. Kvist, and S. Akesson. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. Molecular Ecology 15: 161-171.

Besansky, N. J., D. W. Severson, and M. T. Ferdig. 2003. DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you. Trends in Parasitology 19: 545-546.

Bethenod, M.-T., Y. Thomas, F. Rousset, B. Frerot, L. Pelozuelo, G. Genestier, and D. Bourguet. 2005. Genetic isolation between two sympatric host plant races of the European corn borer, *Ostrinia nubilalis* Hubner. II: assortative mating and host-plant preferences for oviposition. Heredity 94: 264-270.

Bettles, C. M., M. F. Docker, B. Dufour, and D. D. Heath. 2005. Hybridization dynamics between sympartic species of trout: loss of reproductive isolation. Journal of Evolutionary Biology 18: 1220-1233.

Blair, C. P., W. G. Abrahamson, J. A. Jackman, and L. Tyrrell. 2005. Cryptic speciation and host-race formation in a purportedly generalist tumbling flower beetle. Evolution 59: 304-316.

Braendle, C., I. Friebe, M. C. Caillaud, and D. L. Stern. 2005. Genetic variation for an aphid wing polyphenism is genetically linked to a naturally occurring wing polymorphism. Proceedings of the Royal Society of London Series B 272: 657-664.

Brisson, J. A., C. De Toni, I. Duncan, and A. R. Templeton. 2005. Abdominal pigmentation variation in *Drosophila polymorpha*: geographic variation in the trait, and underlying phylogeography. Evolution 59: 1046-1059.

Cardoso, A., and A. P. Vogler. 2005. DNA taxonomy, phylogeny and Pleistocene diversification of the *Cicindela hybrida* species group (Coleoptera: Cincindelidae). Molecular Ecology 14: 3531-3546.

Dayrat, B. 2005. Towards integrative taxonomy. Biological Journal of the Linnean Society 85: 407-415.

Diegisser, T., J. Johannesen, C. Lehr, and A. Seitz. 2004. Genetic and morphological differentiation in *Tephritis bardanae* (Diptera: Tephritidae): evidence for host-race formation. Journal of Evolutionary Biology 17: 83-93.

Donnelly, M. J., J. Pinto, R. Girod, N. J. Besansky, and T. Lehmann. 2004. Revisiting the role of introgression vs. shared ancestral polymorphisms as key processes shaping genetic diversity in the recently separated sibling species of the *Anopheles gambiae* complex. Heredity 92: 61-68.

Dowling, T. E., and C. L. Secor. 1997. The role of hybridization and introgression in the diversification of animals. Annual Review of Ecology and Systematics 28: 593-619.

Du, Y., A. D. Roe, and F. A. H. Sperling. 2005. Phylogenetic framework for *Dioryctria* (Lepidoptera: Pyralidae: Phycitinae) based on combined analysis of mitochondrial DNA and morphology. The Canadian Entomologist 137: 685-711.

Eberhard, W. G. 1985. Sexual Selection and Animal Genitalia. Harvard University Press, Cambridge, MA.

Edwards, S. V., and P. Beerli. 2000. Gene divergence, population divergence and the variance in coalescence time in phylogeographic studies. Evolution 54: 1839-1854.

Emelianov, I., F. Marec, and J. Mallet. 2003. Genomic evidence for divergence with gene flow in host races of the larch budmoth. Proceedings of the Royal Society of London Series B 652: 1-9.

Funk, D. J. 1999. Molecular systematics of cytochrome oxidase I and 16S from *Neochlamisus* leaf beetles and the importance of sampling. Molecular Biology and Evolution 16: 67-82.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution and Systematics 34: 397-423.

Greene, H. W. 1997. Systematics and natural history, foundations for understanding and conserving biodiversity. American Zoologist 34: 48-56.

Gubitz, T., R. S. Thrope, and A. Malhotra. 2005. The dynamics of genetic and morphological variation on volcanic islands. Proceedings of the Royal Society of London Series B 272: 751-757.

Harrison, R. G. 1998. Linking evolutionary pattern and process: the relevance of species concepts for the study of speciation, pp. 19-31. *In* D. J. Howard and S. H. Berlocher [eds.], Endless Forms: Species and Speciation. Oxford University Press, Oxford.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological identification through DNA barcodes. Proceedings of the Royal Society of London Series B 270: 313-321.

Hebert, P. D. N., E. H. Penton, J. M. Burns, D. H. Janzen, and W. Hallwachs. 2005. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences of the United States of America 101: 14812-14817.

Hendrixson, B. E., and J. E. Bond. 2005. Testing species boundaries in the *Anthrodiaetus unicolour* complex (Araneae: Mygalomorphae: Antrodiaetidae): "Paraphyly" and cryptic diversity. Molecular Phylogenetics and Evolution 36: 405-416.

Kiefer, A., F. Mayer, J. Kosuch, O. v. Helversen, and M. Veith. 2002. Conflicting molecular phylogenies of European long-eared bats (*Plecotus*) can be explained by cryptic diversity. Molecular Phylogenetics and Evolution 25: 557-566.

Lipscomb, D., N. Platnick, and Q. Wheeler. 2003. The intellectual content of taxonomy: a comment on DNA taxonomy. Trends in Ecology and Evolution 18: 65-66.

Maddison, W. P. 1997. Gene trees in species trees. Systematic Biology 46: 523-536.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in Ecology & Evolution 20: 229-237.

Mallet, J., and K. Willmot. 2003. Taxonomy: Renaissance or Tower of Babel? Trends in Ecology and Evolution 18: 57-59.

Marko, P. B. 2005. An intraspecific comparative analysis of character divergence between sympatric species. Evolution 59: 554-564.

Meyer, C. P., and G. Paulay. 2005. DNA barcoding: error rates based on comprehensive sampling. Public Library of Science, Biology 3: e422.

Meyerson, L. A., and J. K. Reaser. 2002. Biosecurity: Moving toward a comprehensive approach. Bioscience 52: 593-600.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. Evolution 49: 718-726.

Morando, M., L. Avila, and J. W. J. Sites. 2003. Sampling strategies for delimiting species: genes, individuals, and populations in the *Liolaemus elongatus-kriegi* complex (Squamata: Liolaemidae) in Andean-Patagonia South America. Systematic Biology 52: 1599-185.

Moritz, C., and C. Cicero. 2004. DNA barcoding: promise and pitfalls. Public Library of Science Biology 2: e354.

Mutanen, M. 2005. Delimitation difficulties in species splits: a morphometric case study on the *Euxoa tritici* complex (Lepidoptera, Noctuidae). Systematic Entomology 30: 632-643.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology and Evolution 16: 358-364.

Olsson, U., P. Alstrom, P. G. P. Ericson, and P. Sundberg. 2005. Non-monophyletic taxa and cryptic species - Evidence from a molecular phylogeny of leaf-warblers (*Phylloscopus,* Aves). Molecular Phylogenetics and Evolution 36: 261-276.

Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5: 568-583.

Paquin, P., and M. Hedin. 2004. The power and perils of 'molecular taxonomy': a case study of the eyeless and endangered *Cicurina* (Araneae: Dictynidae) from Texas caves. Molecular Ecology 13: 3239-3255.

Price, T. D., A. Qvarnström, and D. E. Irwin. 2003. The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society of London Series B 270: 1433-1440.

Salzburger, W., S. Baric, and C. Sturmbauer. 2002. Speciation via introgressive hybridization in East African cichlids? Molecular Ecology 11: 619-625.

Sanders, K. L., A. Malhotra, and R. S. Thorpe. 2006. Combining molecular, morphological and ecological data to infer species boundaries in a cryptic tropical pitviper. Biological Journal of the Linnean Society 87: 343-364.

Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proceedings of the National Academy of Sciences of the United States of America 99: 16122-16127.

Sites, J. W., and J. C. Marshall. 2004. Operational criteria for delimiting species. Annual Review of Ecology Evolution and Systematics 35: 199-227.

Strand, M., and P. Sundberg. 2005. Delimiting species in the hoplonemertean genus *Tetrastemma* (Phylum Nemertea): morphology is not concordant with phylogeny as evidenced from mtDNA sequences. Biological Journal of the Linnean Society 86: 201-212.

Takahata, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. Genetics 122: 957-966.

Thalmann, O., D. Serre, M. Hofreiter, D. Lukas, J. Eriksson, and L. Vigilant. 2005. Nuclear insertions help and hinder inference of the evolutionary history of gorilla mtDNA. Molecular Ecology 14: 179-188.

Wahlberg, N., R. Oliveira, and J. A. Scott. 2003. Phylogenetic relationships of *Phyciodes* butterfly species (Lepidoptera: Nymphalidae): complex mtDNA variation and species delimitations. Systematic Entomology 28: 257-273.

Wiens, J. J., and T. A. Penkrot. 2002. Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). Systematic Biology 51: 69-91.

Will, K. W., and D. Rubinoff. 2004. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20: 47-55.

Zhang, D.-X., and G. M. Hewitt. 1996. Nuclear integrations: challenges for mitochondrial DNA markers. Trends in Ecology and Evolution 11: 247-251.

Appendix 1

Specimen information for all examined individuals.

	Species	DNA# (Haplotype)	Abbr	Location	GenBank	Reference
Diptera						
Culicidae	*Anopheles arabiensis* Patton	EM1	*Aa*	Burkina Faso: Zaghtouli: CDC Atlanta, strain ARZAG	DQ792576	Unpublished
Culicidae	*A. gambiae* Giles	-	*Ag*	Strain G3	NC_002084	Beard et al. 1993
Culicidae	*A. gambiae* Giles	EM3 (G1)	*Ag*	Kenya: Asembo Bay: CDC Atlanta, strain AS46	DQ792577	Unpublished
Culicidae	*A. gambiae* Giles	JA37 (G2)	*Ag*	Kenya: Asembo Bay: CDC Atlanta, strain AS46	DQ792578	Unpublished
Culicidae	*A. melas* Theobald	JA22 (ML1)	Am	Gambiae: Balingho: CDC Atlanta, strain BAL	DQ792579	Unpublished
Culicidae	*A. melas* Theobald	EM5 (ML2)	*Am*	Gambiae: Balingho: CDC Atlanta, strain BAL	DQ792580	Unpublished
Culicidae	*A. quadriannulatis* (Theobald)	EM2 (Q1)	*Aq*	Zimbabwe: CDC Atlanta, strain CHIL	DQ792581	Unpublished
Calliphoridae	*Chrysomya albiceps* (Wiedemann)	C_alb5	*Ca*	Egypt: Alexandria: Moharrem Bey	AF083657	(Wells and Sperling 1999)
Calliphoridae	*C. bezziana* Villeneuve	PNG#10	*Cb*	Indonesia: Bogor	AF295548	(Wells and Sperling 2001)
Calliphoridae	*C. megacephala* (F.)	Variant 3	*Cm*	Malaysia	AY909053	Tan et al. 2005
Calliphoridae	*C. megacephala* (F.)	-	*Cm*	Papua New Guinea: btwn Lae & Bulolo	AF295551	Wells and Sperling 2001
Calliphoridae	*C. norrisi James*	PNG#50	*Cn*	Paupa New Guniea: Wau	AF295552	Wells and Sperling 2001
Calliphoridae	*C. rufifacies* (Maquart)	Variant 5	*Cr*	Malaysia	AF909055	Tan et al. 2005
Calliphoridae	*C. rufifacies* (Maquart)	C_ruf2	*Cr*	USA: FL: Miami	AF083658	Wells and Sperling 1999
Calliphoridae	*C. varipes* (Macquart)	C_var1	*Cv*	Australia: Adelaide	AF295556	Wells and Sperling 2001
Lepidoptera						
Tortricidae	*Choristoneura biennis* Freeman	FS.b-53 (b1)	*Chα*	CAN: BC: Morrisey Creek	DQ792587[b]	Sperling and Hickey 1994
Tortricidae	*C. biennis* Freeman	FS.b-54 (bβ)	*Chβ*	CAN: BC: Morrisey Creek	L19096[a]	Sperling and Hickey 1994
Tortricidae	*C. fumiferana* (Clemens)	FS.b-12 (f2)	*Chf*	CAN: ON: Ignace	L19098[a]	Sperling and Hickey 1994
Tortricidae	*C. fumiferana* (Clemens)	FS.b-37 (f1)	*Chf*	CAN: AB: Manning	L19094[a]	Sperling and Hickey 1994
Tortricidae	*C. occidentalis* Freeman	FS.b-16 (o1)	*Chα*	CAN: BC: Monte Creek	DQ792584[b]	Sperling and Hickey 1994
Tortricidae	*C. occidentalis* Freeman	FS.b-367 (oβ)	*Chβ*	CAN: BC: Bridesville	DQ792585[b]	Sperling and Hickey 1994
Tortricidae	*C. orae* Freeman	FS.b-216 (b1)	*Chα*	USA: AK: Fairbanks; phero. 82:9:9	DQ792586[b]	Sperling and Hickey 1994
Tortricidae	*C. pinus* (Freeman)	FS.b-15 (p1)	*Chα*	CAN: ON: Parry Sound	L19095[a]	Sperling and Hickey 1994
Tortricidae	*C. retiniana* (Walsingham)	FS.b-816	*Chβ*	USA: CA: Sierraville	DQ792588	Unpublished
Tortricidae	*C. retiniana* (Walsingham)	FS.b-817	*Chα*	USA: CA: Tehachapi	DQ792590	Unpublished
Tortricidae	*C. retiniana* (Walsingham)	FS.b-866	*Chβ*	USA: NV: Mt. Charleston	DQ792589	Unpublished
Pyralidae	*Dioryctria abietella* Denis and Schiffermüller	Du64	*Da*	China: Henan Province: Mt. Baiyun	DQ247739	Du et al. 2005
Pyralidae	*D. abietivorella* (Grote)	AR22	*Dav*	USA: CA: Butte Co.: Chico	DQ295185	Roe et al. 2006

	Species	DNA# (Haplotype)	Abbr	Location	GenBank	Reference
Pyralidae	*D. abietivorella* (Grote)	Du04	*Dav*	USA: CA: Butte Co.: Chico	DQ247740	Du et al. 2005
Pyralidae	*D. abietivorella* (Grote)	Du05	*Dav*	USA: CA: Butte Co.: Chico	DQ247741	Du et al. 2005
Pyralidae	*D. auranticella* (Grote)	AR144	*Dau*	USA: CA: El Dorado Co., Placerville	DQ295176	Roe *et al.* 2006
Pyralidae	*D. auranticella* (Grote)	Du02	*Dau*	USA: CA: El Dorado Co., Placerville	DQ247736	Du et al. 2005
Pyralidae	*D. cambiicola* (Grote)	AR78	*Dc*	CAN: BC: Prince George	DQ295183	Roe *et al.* 2006
Pyralidae	*D. fordi* Donahue and Neunzig	AR157	*Df*	USA: CA: Butte Co. Chico	DQ295184	Roe *et al.* 2006
Pyralidae	*D. magnifica* Munroe	Du69	*Dm*	China: Mt. Baiyun, Henan Province	DQ247742	Du et al. 2005
Pyralidae	*D. okanaganella* Mutuura, Munroe & Ross	AR150	*Do*	USA: CA: El Dorado Co., Placerville	DQ295179	Roe *et al.* 2006
Pyralidae	*D. okanaganella* Mutuura, Munroe & Ross	AR148	*Do*	USA: CA: El Dorado Co., Placerville	DQ295178	Roe *et al.* 2006
Pyralidae	*D. pentictonella* Mutuura, Munroe & Ross	AR15	*Dp*	USA: CA: Butte Co.: Chico	DQ295180	Roe *et al.* 2006
Pyralidae	*D. pentictonella* Mutuura, Munroe & Ross	AR58	*Dp*	USA: CA: Butte Co.: Chico	DQ295181	Roe *et al.* 2006
Pyralidae	*D. pentictonella* Mutuura, Munroe & Ross	AR149	*Dp*	USA: CA: El Dorado Co., Placerville	DQ295182	Roe *et al.* 2006
Pyralidae	*D. pseudotsugella* Munroe	AR82	*Dps*	CAN: AB: Vulcan: 10 mi SE	DQ295186	Roe *et al.* 2006
Pyralidae	*D. reniculelloides* Mutuura and Munroe	Du01	*Dre*	CAN: AB: Fort MacMurray	DQ247734	Du et al. 2005
Pyralidae	*D. rubella* Hampson	Du21	*Dr*	China: Tianjin: Mt. Baxian	DQ247743	Du et al. 2005
Pyralidae	*D. sylvestrella* (Ratzeburg)	D130	*Ds*	Germany: Bavaria: Parkstein-Hütten	DQ247746	Du et al. 2005
Pyralidae	*D. sylvestrella* (Ratzeburg)	02087	*Ds*	Germamy: Bavaria: Landshut	DQ247745	Du et al. 2005
Pyralidae	*D. yiai* Mutuura and Munroe	Du13	*Dy*	China: Hebei Province: Mt. Xiaowutai	DQ247737	Du et al. 2005
Pyralidae	*D. yiai* Mutuura and Munroe	D17	*Dy*	China: Tianjin: Mt. Baxian	DQ247738	Du et al. 2005
Pyralidae	*D. zimmermani* (Grote)	Du118	*Dz*	USA: MS: Hinds Co.	DQ247730	Du et al. 2005
Noctuidae	*Feltia jaculifera* (Gn.)	FS.b-150 (pheroA)	*Fj*	CAN: AB: Lethbridge	U60990	Sperling et al. 1996
Noctuidae	*F. jaculifera* (Gn.)	FS.b-152 (pheroB)	*Fj*	CAN: AB: Lethbridge	DQ792591[b]	Sperling et al. 1996
Noctuidae	*F. jaculifera* (Gn.)	FS.b-464 (pheroC)	*Fj*	CAN: ON: Ottawa	DQ792592[b]	Sperling et al. 1996
Sphingidae	*H. annei* Guérin	16157	*Ha*	Chile: Toconao, north	AJ749430	Hundsoerfer et al. 2005
Sphingidae	*H. e. euphorbiae* (L.)	695887	*He*	France: South	AJ749480	Hundsoerfer et al. 2005
Sphingidae	*H. e. euphorbiae* (L.)	020c	*He*	Spain: Alió: Catalonia	AJ749485	Hundsoerfer et al. 2005
Sphingidae	*H. e. euphorbiae* (L.)	23172a	*He*	Germany	AJ749512	Hundsoerfer et al. 2005
Sphingidae	*H. e. euphorbiae* (L.)	010	*He*	Spain	AJ749514	Hundsoerfer et al. 2005
Sphingidae	*H. euphorbiarum* (Guérin-Mereville & Percheron)	23274	*Hem*	Argentina	AJ749428	Hundsoerfer et al. 2005
Sphingidae	*H. gallii* (Rottemburg)	0042	*Hg*	China: Yanqin: near Beijing	AJ749432	Hundsoerfer et al. 2005
Sphingidae	*H. gallii* (Rottemburg)	695869	*Hg*	England: ex bred stock, ex Neil West	AJ749433	Hundsoerfer et al. 2005
Sphingidae	*H. gallii* (Rottemburg)	0027	*Hg*	Czech Republic: ex bred stock	AJ749450	Hundsoerfer et al. 2005
Sphingidae	*H. gallii* (Rottemburg)	16189	*Hg*	Germany	AJ749451	Hundsoerfer et al. 2005
Sphingidae	*H. gallii* (Rottemburg)	0080	*Hg*	China	AJ749579	Hundsoerfer et al. 2005
Sphingidae	*H. gallii* (Rottemburg)	695843	*Hg*	Finland	AJ749580	Hundsoerfer et al. 2005
Sphingidae	*H. nicaea castissima* (Austaut)	23208	*Hn*	Morocco: High Atlas, SW Midelt	AJ749444	Hundsoerfer et al. 2005

	Species	DNA# (Haplotype)	Abbr	Location	GenBank	Reference
Sphingidae	*H. robertsi peplidis* (Christoph)	695835	*Hrp*	Iran: near Esfahan	AJ749464	Hundsoerfer et al. 2005
Sphingidae	*H. r. peplidis* (Christoph)	695842	*Hrp*	Iran: near Esfahan	AJ749465	Hundsoerfer et al. 2005
Sphingidae	*H. r. peplidis* (Christoph)	695872	*Hrp*	Iran: near Esfahan	AJ749466	Hundsoerfer et al. 2005
Sphingidae	*H. sammuti* Eitchberger, Danner & Surholt	0053	*Hs*	Malta	AJ749505	Hundsoerfer et al. 2005
Sphingidae	*H. sammuti* Eitchberger, Danner & Surholt	23239	*Hs*	Italy: Sicily: Zafferana	AJ749459	Hundsoerfer et al. 2005
Sphingidae	*H. sammuti* Eitchberger, Danner & Surholt	0054	*Hs*	Malta	AJ749461	Hundsoerfer et al. 2005
Sphingidae	*H. sammuti* Eitchberger, Danner & Surholt	0065	*Hs*	Malta	AJ749463	Hundsoerfer et al. 2005
Sphingidae	*H. tithymali deserticola* (Staudinger)	055d	*Ht*	Morocco	AJ749497	Hundsoerfer et al. 2005
Sphingidae	*H. t. deserticola* (Staudinger)	055c	*Ht*	Morocco	AJ749496	Hundsoerfer et al. 2005
Sphingidae	*H. t. deserticola* (Staudinger)	055b	*Ht*	Morocco	AJ749495	Hundsoerfer et al. 2005
Sphingidae	*H. t. himyarensis* Meerman	23173a	*Ht*	Yemen	AJ749521	Hundsoerfer et al. 2005
Sphingidae	*H. t. himyarensis* Meerman	23224	*Ht*	Yemen	AJ749499	Hundsoerfer et al. 2005
Sphingidae	*H. t. mauretanica* (Staudinger)	23215	*Ht*	Morocco	AJ749494	Hundsoerfer et al. 2005
Sphingidae	*H. t. mauretanica* (Staudinger)	23216	*Ht*	Morocco	AJ749545	Hundsoerfer et al. 2005
Sphingidae	*H. t. tithymali* (Boisduval)	084a	*Ht*	Spain: Canary Islands	AJ479486	Hundsoerfer et al. 2005
Sphingidae	*H. t. tithymali* (Boisduval)	100_AH	*Ht*	Spain: Canary Islands	AJ479488	Hundsoerfer et al. 2005
Sphingidae	*H. t. gecki* de Freina	23238	*Ht*	Portugal	AJ749491	Hundsoerfer et al. 2005
Sphingidae	*H. t. gecki* de Freina	0164	*Ht*	Portugal	AJ749490	Hundsoerfer et al. 2005
Geometridae	*L. fiscellaria fiscellaria* (Guenée)	FS.b-#9 (F1)	*Lf*	Canada: NF: Corner Brook	AF064521	Sperling et al. 1999
Geometridae	*L. f. lugubrosa* (Hulst)	FS.b-265 (L1)	*Lf*	Canada: BC: Mud Lake	DQ792593[b]	Sperling et al. 1999
Papilionidae	*Papilio anchisiades* Esper	FS.a-44	*Pa*	Brazil: Campinas	AF044005	Caterino and Sperling 1999
Papilionidae	*P. canadensis* Rothschild & Jordan	FS.a-16	*Pca*	USA: NY: Richford	AF044014	Caterino and Sperling 1999
Papilionidae	*P. cresphontes* Cramer	FS.a-170	*Pc*	USA: WI: Sauk County	AF043999	Caterino and Sperling 1999
Papilionidae	*P. demodocus demodocus* Esper	FS.b-1929	*Pd*	South Africa: Nelspruit	AY569091	Zakharov et al. 2004b
Papilionidae	*P. demodocus* Esper	FS.b-146	*Pd*	Kenya (ex pupa)	AY457588	Zakharov et al. 2004a
Papilionidae	*P. demoleus malaynus*	FS.a-68	*Pde*	Malaysia: Penang Island	AF044000	Caterino and Sperling 1999
Papilionidae	*P. d. sthenelus* W.S. Macleay	FS.b-1832	*Pde*	Australia: New South Wales	AY569092	Zakharov et al. 2004b
Papilionidae	*P. erithonioides* Grose-Smith	FS.b-1626	*Pe*	Madagascar: Ankitsanga	AY565095	Zakharov et al. 2004b
Papilionidae	*P. erostratus* Westwood	FS.b-973	*Per*	El Salvador	AY457599	Zakharov et al. 2004a
Papilionidae	*P. glaucus* L.	FS.a-69	*Pgl*	USA: MD: Potomac	AF044013	Caterino and Sperling 1999
Papilionidae	*P. grosesmithi* Rothschild	FS.b-1625	*Pg*	Madagascar: Kirindy	AY569089	Zakharov et al. 2004b
Papilionidae	*P. grosesmithi* Rothschild	FS.b-1624	*Pg*	Madagascar: Ambahibe	AY569090	Zakharov et al. 2004b
Papilionidae	*P. hospiton* Géné	FS.a-143	*Ph*	Sardinia	AF044009	Caterino and Sperling 1999
Papilionidae	*P. indra* Reakirt	FS.a-66	*Pi*	USA: WA: Wawawai	AF044011	Caterino and Sperling 1999
Papilionidae	*P. machaon machaon* L.	FS.a-27	*Pma*	France: Coudoux	AF044006	Caterino and Sperling 1999

	Species	DNA# (Haplotype)	Abbr	Location	GenBank	Reference
Papilionidae	*P. m. hippocrates* Felder x Felder	FS.b-78	*Pmh*	Japan: Gifu Pref.	AY457593	Zakharov et al. 2004a
Papilionidae	*P. m. oregonius* Edwards	FS.a-77	*Pmo*	USA: WA: Palouse Falls	AF044007	Caterino and Sperling 1999
Papilionidae	*P. memnon* L.	FS.b-91	*Pm*	Japan: Gifu Pref.	AY457578	Zakharov et al. 2004a
Papilionidae	*P. multicaudatus* Kirby	FS.a-163	*Pmu*	USA: SD: Black Hills	AF044016	Caterino and Sperling 1999
Papilionidae	*P. palamedes* Drury	FS.a-18	*Pp*	USA: FL: Ocala State Forest	AF044018	Caterino and Sperling 1999
Papilionidae	*P. polyxenes* Fabricius	FS.a-64	*Ppo*	USA: NY: Tompkins County	AF044010	Caterino and Sperling 1999
Papilionidae	*P. rumanzovia* Eshscholtz	FS.b-972	*Pr*	Phillippines	AY457582	Zakharov et al. 2004a
Papilionidae	*P. rutulus* Lucas	FS.a-472	*Pru*	USA: WA: Orcas Island	AF044015	Caterino and Sperling 1999
Papilionidae	*P. scamander* Boisduval	FS.a-19	*Ps*	Brazil: Campinas	AF044020	Caterino and Sperling 1999
Papilionidae	*P. thoas* L.	FS.b-302	*Pth*	French Guiana: Pointe Macouria	AY457601	Zakharov et al. 2004a
Papilionidae	*P. troilus* L.	FS.a-29	*Pt*	USA: FL: Ocala State Forest	AF044017	Caterino and Sperling 1999
Papilionidae	*P. xuthus* L.	FS.a-238	*Px*	Japan: Tokyo	AF043999	Caterino and Sperling 1999
Papilionidae	*P. zelicaon* Lucas	FS.a-76	*Pz*	USA: CA: Riverside County	AF044008	Caterino and Sperling 1999
Yponomeutidae	*Yponomeuta cagnagella* Hubner	FS.b-467	*Yc*	Canada: ON: Ottawa	DQ792583[b]	Sperling et al. 1995
Yponomeutidae	*Y. padella* (L.)	FS.b-470	*Yp*	Canada: BC: Victoria.	DQ792582[b]	Sperling et al. 1995

[a] partial COI-COII fragment extended to full 2.3 kb COI-COII
[b] previously published but not submitted to GenBank

Appendix 2

Collection information and voucher numbers for all specimens examined for molecular and morphometric analysis.

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
Canada	Alberta	Belly River Campground, Waterton Lakes National Park	49.025	-113.680	*Picea glauca*	M	A. Roe and L. Lumley	8-Jul-05	AR361
Canada	Alberta	Dom. Ranger Sta. Manyberries	49.400	-110.700	unknown	F	D. F. Hardwick	23-Jul-55	UASM 3290
Canada	Alberta	Cypress Hills	49.667	-109.500	*Picea glauca*	M	L. Lumley	10-Jun-05	AR368
Canada	Alberta	Elkwater	49.667	-110.283	*Picea glauca*	F	-	3-Aug-55	NFRC 100106
Canada	Alberta	Elkwater	49.667	-110.283	*Picea glauca*	M	-	29-Jul-55	NFRC 100107
Canada	Alberta	Elkwater	49.667	-110.283	*Picea glauca*	F	-	27-Jul-59	NFRC 100149
Canada	Alberta	Elkwater	49.667	-110.283	*Picea glauca*	F	-	20-Jul-59	NFRC 100155
Canada	Alberta	Elkwater, 15mi E	49.667	-110.283	*Picea glauca*	F	-	20-Jul-59	NFRC 100160
Canada	Alberta	Elkwater	49.667	-110.283	*Picea glauca*	F	-	25-Jul-55	NFRC 100168
Canada	Alberta	Elkwater	49.667	-110.283	*Picea glauca*	F	-	14-Aug-55	NFRC 100169
Canada	Alberta	Lethbridge	49.700	-112.817	unknown	M	E. E. Sterns	13-Jul-60	UASM 3293
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	AR273
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	AR274
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	AR275
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	AR276
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	AR277
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	F	A. Roe	15-Aug-02	AR278
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	AR279
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	AR280
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D31
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D32
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D33
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	F	A. Roe	15-Aug-02	D34
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D35
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D36
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D37
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D38
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D39
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D40
Canada	Alberta	Beaver Creek Campground, ~30km west Ft. McLeod	49.804	-113.935	unknown	M	A. Roe	15-Aug-02	D41

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
Canada	Alberta	E Porcupine Hills	49.845	-114.264	*Pseudotsuga menziesii*	F	A. Roe and L. Lumley	6-Jul-05	AR360
Canada	Alberta	Nobleford	49.883	-113.050	*Picea glauca*	-	-	-	UASM 3278
Canada	Alberta	Dutch Creek Campgroud	49.909	-114.391	unknown	M	A. Roe	5-Aug-01	AR210
Canada	Alberta	Walsh Campground	49.950	-110.050	*Picea glauca*	F	FIDS	-	UASM 3276
Canada	Alberta	Walsh Campground	49.950	-110.050	*Picea glauca*	F	FIDS	-	UASM 3289
Canada	Alberta	Walsh Campground	49.950	-110.050	*Picea glauca*	F	FIDS	-	UASM 3292
Canada	Alberta	Vulcan, 7mi E and 3 mi S, along Lomond Rd.	50.358	-113.096	unknown	M	A. Roe	7-Aug-01	AR208
Canada	Alberta	Vulcan, 7mi E and 3 mi S, along Lomond Rd.	50.358	-113.096	unknown	F	A. Roe	7-Aug-01	AR82
Canada	Alberta	Vulcan, 7mi E and 3 mi S, along Lomond Rd.	50.358	-113.096	unknown	F	A. Roe	7-Aug-01	AR83
Canada	Alberta	Vulcan, 7mi E and 3 mi S, along Lomond Rd.	50.358	-113.096	unknown	F	A. Roe	7-Aug-01	D09
Canada	Alberta	Vulcan, 7mi E and 3 mi S, along Lomond Rd.	50.358	-113.096	unknown	F	A. Roe	7-Aug-01	D10
Canada	Alberta	Vulcan, 7mi E and 3 mi S, along Lomond Rd.	50.358	-113.096	unknown	F	A. Roe	7-Aug-01	D11
Canada	Alberta	Vulcan, 7mi E and 3 mi S, along Lomond Rd.	50.358	-113.096	unknown	F	A. Roe	7-Aug-01	D12
Canada	Alberta	8 km SSE Seebe, Kananaskis Research Stn.	51.100	-115.067	unknown	M	B. C. Schmidt	1-Aug-05	UASM 3589
Canada	Alberta	Tolman Bridge	51.834	-113.008	unknown	-	J.-F. Landry	23-Jul-03	AR313
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	F	A. Roe	24-Jul-03	AR309
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	F	A. Roe	24-Jul-03	AR311
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	F	A. Roe	24-Jul-03	D01
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	F	A. Roe	24-Jul-03	D02
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	M	A. Roe	24-Jul-03	D03
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	F	A. Roe	24-Jul-03	D04
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	M	A. Roe	24-Jul-03	D05
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	F	A. Roe	24-Jul-03	D06
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	-	A. Roe	24-Jul-03	D07
Canada	Alberta	Bearberry, 7 mi E along R.R. 7.20	51.839	-114.710	unknown	F	A. Roe	24-Jul-03	D08
Canada	Alberta	52°N 115°W	52.000	-115.000	*Picea glauca*	M	-	20-Jul-55	NFRC 100150
Canada	Alberta	Thompson Cr. Rec Area	52.012	-116.628	*Picea glauca*	F	L. Lumley	10-Jun-05	AR369
Canada	Alberta	Thompson Cr. Rec Area	52.012	-116.628	*Picea glauca*	M	L. Lumley	10-Jun-05	AR370
Canada	Alberta	Thompson Cr. Rec Area	52.012	-116.628	*Picea glauca*	M	L. Lumley	10-Jun-05	AR371
Canada	Alberta	Thompson Cr. Rec Area	52.012	-116.628	*Picea glauca*	-	L. Lumley	10-Jun-05	AR372
Canada	Alberta	Innisfail	52.033	-113.950	*Picea glauca*	M	-	30-Jul-54	NFRC 100161
Canada	Alberta	Provost	52.350	-110.267	*Picea glauca*	M	-	26-Jul-54	NFRC 100159
Canada	Alberta	Provost	52.350	-110.267	*Picea pungens*	F	-	12-Jul-51	NFRC 100163
Canada	Alberta	Edmonton, Windsor Park area	53.550	-113.467	unknown	F	F. Sperling	8-Jul-02	UASM 3612

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
Canada	Alberta	Spedden	54.133	-111.717	*Picea glauca*	F	-	20-Jul-54	NFRC 100158
Canada	Alberta	Seibert Lake	54.692	-111.264	unknown	F	L. Lumley	27-Jul-05	AR357
Canada	Alberta	Seibert Lake	54.692	-111.264	unknown	F	L. Lumley	27-Jul-05	AR358
Canada	Alberta	Smith	55.167	-114.033	*Picea glauca*	M	-	14-Jul-53	NFRC 100111
Canada	Alberta	Hythe	55.331	-119.455	*Picea glauca*	F	L. Lumley	13-Jun-05	AR359
Canada	Alberta	Dunvegan	55.917	-118.600	*Picea glauca*	M	-	10-Jul-52	NFRC 100171
Canada	Alberta	Peace River	56.250	-117.283	*Picea glauca*	M	-	28-Jun-65	NFRC 100145
Canada	Alberta	Warrensville	56.300	-117.667	*Picea glauca*	M	-	13-Jul-53	NFRC 100151
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	AR203
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	F	A. Roe	11-Jul-01	AR205
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	AR206
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	AR207
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	F	A. Roe	11-Jul-01	AR73
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	AR74
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	AR75
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	AR77
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D15
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D16
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D18
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D19
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D20
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D21
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D22
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	?	A. Roe	11-Jul-01	D23
Canada	Alberta	Fort MacMurray	56.376	-111.378	unknown	M	A. Roe	11-Jul-01	D24
Canada	Alberta	Lone Star	56.717	-117.650	*Picea glauca*	F	-	20-Jul-54	NFRC 100152
Canada	Alberta	Lone Star	56.717	-117.650	*Picea glauca*	F	-	20-Jul-54	NFRC 100157
Canada	Alberta	Lone Star	56.717	-117.650	*Picea glauca*	M	-	17-Jul-54	NFRC 100215
Canada	Alberta	Hotchkiss	57.067	-117.550	*Picea glauca*	F	-	23-Aug-53	NFRC 100104
Canada	Alberta	Hotchkiss	57.067	-117.550	*Picea glauca*	F	-	16-Jul-53	NFRC 100164
Canada	Alberta	Hotchkiss	57.067	-117.550	*Picea glauca*	F	-	13-Jul-53	NFRC 100165
Canada	Alberta	Rainbow Lake	58.297	-119.404	*Picea glauca*	F	L. Lumley	15-Jun-05	AR367
Canada	Alberta	Upper Hay River	59.056	-117.758	*Picea glauca*	F	-	14-Aug-53	NFRC 100105
Canada	Alberta	La Butte Provincial Park	59.376	-111.201	unknown	M	A. Roe	10-Jul-01	AR209

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
Canada	Alberta	La Butte Provincial Park	59.376	-111.201	unknown	M	A. Roe	10-Jul-01	AR72
Canada	Alberta	Fort Smith	60.000	-111.850	*Picea glauca*	F	-	15-Jul-52	NFRC 100173
Canada	Alberta	P.P. 8-028-249	-	-	*Picea glauca*	F	-	23-Jul-55	NFRC 100109
Canada	Alberta	P.P. 8-028-249	-	-	*Picea glauca*	F	-	25-Jul-55	NFRC 100110
Canada	Alberta	Remple Mill Rd.	-	-	*Picea glauca*	M	-	19-Jul-54	NFRC 100162
Canada	Alberta	P.P. 8-028-249	-	-	*Picea glauca*	M	-	26-Jun-55	NFRC 100166
Canada	Alberta	P.P. 8-028-249	-	-	*Picea glauca*	F	-	25-Jul-55	NFRC 100172
Canada	Alberta	P.P. 8-028-249	-	-	*Picea glauca*	F	-	26-Jul-56	NFRC 100174
Canada	Alberta	Chendle	-	-	*Picea pungens bud*	F	-	25-Jul-55	NFRC 100175
Canada	British Columbia	Elko, 11km SSW	46.264	-115.257	unknown	F	C. Schmidt	-	UASM 3277
Canada	British Columbia	Elko, 11km SSW	46.264	-115.257	unknown	F	C. Schmidt	9-Aug-03	UASM 3557
Canada	British Columbia	Beech Head	48.317	-123.650	*Pseudotsuga menziesii*	M	FIDS	9-Aug-03	UASM 3558
Canada	British Columbia	Lumberton	49.417	-115.867	*Picea engelmanni*	F	FIDS	-	UASM 3323
Canada	British Columbia	10 mi E Cranbrook	49.500	-115.767	unknown	F	D. F. Hardwick	-	UASM 3298
Canada	British Columbia	10 mi E Cranbrook	49.500	-115.767	unknown	F	D. F. Hardwick	29-Jul-64	UASM 3324
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	29-Jul-64	UASM 3325
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D43
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D44
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D45
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D46
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D47
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D48
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D49
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D50
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D51
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D52
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D53
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	F	A. Roe	4-Aug-03	D54
Canada	British Columbia	Summerland, ~30 km west	49.720	-120.178	unknown	M	A. Roe	4-Aug-03	D55
Canada	British Columbia	Ta Ta Cr.	49.783	-115.783	*Picea engelmanni*	F	FIDS	-	UASM 3294
Canada	British Columbia	Vernon: Kalamalka Prov. Park	50.208	-119.269	unknown	F	A. Roe	9-Aug-05	AR363
Canada	British Columbia	Vernon: Kalamalka Prov. Park	50.208	-119.269	unknown	F	A. Roe	9-Aug-05	AR364
Canada	British Columbia	Vernon: Kalamalka Prov. Park	50.208	-119.269	unknown	F	A. Roe	9-Aug-05	AR365
Canada	British Columbia	Vernon: Kalamalka Prov. Park	50.208	-119.269	unknown	M	A. Roe	9-Aug-05	AR366

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
Canada	British Columbia	Fairmont	50.317	-115.867	*Picea engelmanni*	M	FIDS	-	UASM 3299
Canada	British Columbia	Vernon: Silverstar Provincial Park	50.344	-119.115	unknown	F	A. Roe	9-Aug-05	AR362
Canada	British Columbia	Pritchard, 10 km E along Old Duck Rd.	50.599	-119.892	unknown	M	A. Roe	6-Aug-03	D13
Canada	British Columbia	Pritchard, 10 km E along Old Duck Rd.	50.599	-119.892	unknown	F	A. Roe	6-Aug-03	D14
Canada	British Columbia	Grandview Bench	50.650	-119.150	*Pseudotsuga menziesii*	F	FIDS	-	UASM 3322
Canada	British Columbia	Seton Lake, Lillooet	50.683	-122.117	unknown	F	J. McDunnough	14-Jul-30	UASM 3295
Canada	British Columbia	Hat Creek	50.883	-121.400	*Pseudotsuga menziesii*	M	FIDS	-	UASM 3301
Canada	British Columbia	Adams Lake Provincial Park	50.945	-119.668	unknown	-	A. Roe	5-Aug-03	AR314
Canada	British Columbia	Adams Lake Provincial Park	50.945	-119.668	unknown	F	A. Roe	5-Aug-03	D25
Canada	British Columbia	Adams Lake Provincial Park	50.945	-119.668	unknown	F	A. Roe	5-Aug-03	D26
Canada	British Columbia	Adams Lake Provincial Park	50.945	-119.668	unknown	F	A. Roe	5-Aug-03	D27
Canada	British Columbia	Adams Lake Provincial Park	50.945	-119.668	unknown	M	A. Roe	5-Aug-03	D28
Canada	British Columbia	Adams Lake Provincial Park	50.945	-119.668	unknown	F	A. Roe	5-Aug-03	D29
Canada	British Columbia	Adams Lake Provincial Park	50.945	-119.668	unknown	M	A. Roe	5-Aug-03	D30
Canada	British Columbia	Aleza Lake	54.168	-122.033	*Picea glauca*	F	FIDS	-	UASM 3300
Canada	British Columbia	Lac La Jeune	-	-	*Pseudotsuga menziesii*	M	FIDS	-	UASM 3302
Canada	Manitoba	Spruce Woods Forest Reserve	49.777	-99.350	*Picea* sp.	F	-	2-Jul-48	NFRC 100183
Canada	Manitoba	Spruce Woods Forest Reserve	49.777	-99.350	*Picea glauca*	F	-	17-Jul-52	NFRC 100197
Canada	Manitoba	Spruce Woods Forest Reserve	49.777	-99.350	*Picea glauca*	F	-	8-Jul-55	NFRC 100198
Canada	Manitoba	Spruce Woods	49.821	-99.652	*Picea mariana*	M	-	27-Jun-41	NFRC 100114
Canada	Manitoba	Spruce Woods	49.821	-99.652	*Picea glauca*	M	-	3-Jul-46	NFRC 100185
Canada	Manitoba	Portage la Prairie	49.973	-98.292	*Picea pungens*	F	-	24-Jul-67	NFRC 100112
Canada	Manitoba	Oak Point	50.506	-98.029	*Acer negundo*	F	-	9-Jul-57	NFRC 100178
Canada	Manitoba	Gimli	50.634	-96.990	*Picea glauca*	F	-	11-Jul-56	NFRC 100154
Canada	Manitoba	Garner Lake	50.816	-95.184	*Picea glauca*	M	-	13-Jul-56	NFRC 100226
Canada	Manitoba	Arborg	50.907	-97.217	*Picea glauca*	F	-	28-Jul-69	NFRC 100186
Canada	Manitoba	Shevlin	51.194	-101.208	*Picea glauca*	unknown	-	22-Jul-66	NFRC 100115
Canada	Manitoba	Namew Lake	54.237	-101.934	*Picea glauca*	F	-	7-Jul-55	NFRC 100188
Canada	Manitoba	Namew Lake	54.237	-101.934	*Picea glauca*	unknown	-	8-Jul-55	NFRC 100193
Canada	Manitoba	Namew Lake	54.237	-101.934	*Picea glauca*	F	-	8-Jul-55	NFRC 100194
Canada	Manitoba	Cranberry Portage	54.586	-101.377	*Picea glauca*	F	-	24-Jul-64	NFRC 100182
Canada	Manitoba	Cranberry Portage	54.586	-101.377	*Picea glauca*	F	-	22-Jul-64	NFRC 100196
Canada	Manitoba	Mistik Lake	54.629	-101.552	*Picea glauca*	F	-	17-Jul-64	NFRC 100192

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
Canada	Manitoba	Rosenberg	54.801	-101.761	*Picea glauca*	-	-	23-Jun-55	NFRC 100227
Canada	Northwest Territories	Fort Resolution	61.171	-113.671	*Picea glauca*	F	-	22-Jul-65	NFRC 100201
Canada	Northwest Territories	Fort Resolution	61.171	-113.671	*Picea glauca*	F	-	21-Jul-65	NFRC 100202
Canada	Northwest Territories	Fort Resolution	61.171	-113.671	*Picea glauca*	M	-	21-Jul-65	NFRC 100203
Canada	Northwest Territories	Fort Resolution	61.171	-113.671	*Picea glauca*	F	-	21-Jul-65	NFRC 100204
Canada	Northwest Territories	Fort Resolution	61.171	-113.671	*Picea glauca*	M	-	26-Jul-65	NFRC 100205
Canada	Northwest Territories	Fort Resolution	61.171	-113.671	*Picea glauca*	F	-	26-Jul-65	NFRC 100206
Canada	Northwest Territories	Liard River, mouth	61.849	-121.308	*Picea glauca*	F	-	13-Jul-55	NFRC 100207
Canada	Northwest Territories	Liard River, mouth	61.849	-121.308	*Picea glauca*	F	-	8-Jul-55	NFRC 100208
Canada	Northwest Territories	62°N 123°W	62.000	-123.000	*Picea glauca*	M	-	11-Jul-55	NFRC 100213
Canada	Northwest Territories	62°N 123°W	62.000	-123.000	*Picea glauca*	M	-	13-Jul-55	NFRC 100214
Canada	Northwest Territories	McKenzie River 124°46'W 64°27'N	64.450	-124.767	*Picea glauca*	M	-	21-Jul-56	NFRC 100217
Canada	Northwest Territories	McKenzie River 124°46'W 64°27'N	64.450	-124.767	*Picea glauca*	M	-	21-Jul-56	NFRC 100218
Canada	Northwest Territories	McKenzie River 124°46'W 64°27'N	64.450	-124.767	*Picea glauca*	M	-	21-Jul-56	NFRC 100219
Canada	Northwest Territories	Fort Norman	64.901	-125.578	*Picea glauca*	F	-	25-Jul-59	NFRC 100118
Canada	Northwest Territories	Fort Norman	64.901	-125.578	*Picea glauca*	F	-	29-Jul-59	NFRC 100209
Canada	Northwest Territories	Fort Norman	64.901	-125.578	*Picea glauca*	F	-	28-Jul-59	NFRC 100210
Canada	Northwest Territories	Fort Norman	64.901	-125.578	*Picea glauca*	M	-	27-Jul-59	NFRC 100211
Canada	Northwest Territories	Fort Norman	64.901	-125.578	*Picea glauca*	F	-	22-Jul-55	NFRC 100212
Canada	Northwest Territories	Bear Rock, near Fort Norman	64.967	-125.719	*Picea glauca*	F	-	23-Jul-56	NFRC 100222
Canada	Northwest Territories	Bear Rock, near Fort Norman	64.967	-125.719	*Picea glauca*	F	-	23-Jul-56	NFRC 100223
Canada	Northwest Territories	Bear Rock, near Fort Norman	64.967	-125.719	*Picea glauca*	F	-	23-Jul-56	NFRC 100224
Canada	Northwest Territories	Bear Rock, near Fort Norman	64.967	-125.719	*Picea glauca*	F	-	23-Jul-56	NFRC 100225
Canada	Northwest Territories	-	-	-	*Picea glauca*	-	FIDS	-	UASM 3291
Canada	Ontario	Timmins	48.467	-81.333	*Picea mariana*	F	-	6-Jul-42	NFRC 100119
Canada	Ontario	Plays Plat	-	-	unknown	F	-	21-Jul-43	NFRC 100132
Canada	Saskachewan	Cypress Hills	49.667	-109.500	*Picea glauca*	M	-	11-Jul-58	NFRC 100124
Canada	Saskachewan	Cypress Hills	49.667	-109.500	*Picea glauca*	M	-	18-Jul-58	NFRC 100125
Canada	Saskachewan	Cypress Hills	49.667	-109.500	*Picea glauca*	M	-	14-Jul-58	NFRC 100139
Canada	Saskachewan	Indian Head	50.533	-103.667	*Picea glauca*	M	-	5-Jul-57	NFRC 100128
Canada	Saskachewan	Madge Lake	51.667	-101.633	*Populous tremuloides*	F	-	8-Jul-58	NFRC 100113
Canada	Saskachewan	Greenbush Rd.	52.850	-102.700	*Picea glauca*	F	-	11-Jul-56	NFRC 100134
Canada	Saskachewan	MacDowall	53.017	-106.017	*Picea glauca*	F	-	2-Jul-56	NFRC 100179
Canada	Saskachewan	Prince Albert	53.200	-105.767	*Picea glauca*	-	-	14-Jul-53	NFRC 100228

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
Canada	Saskachewan	Candle Lake	53.833	-105.300	*Picea glauca*	M	-	4-Jul-68	NFRC 100135
Canada	Saskachewan	Amisk Lake	54.583	-102.250	*Picea glauca*	F	-	21-Jul-64	NFRC 100143
Canada	Saskachewan	Barkwell	57.617	-106.100	*Picea glauca*	M	-	2-Jul-56	NFRC 100127
Canada	Saskachewan	Viellardville	-	-	*Picea glauca*	M	-	11-Jul-56	NFRC 100133
Canada	Saskachewan	-	-	-	*Picea glauca*	M	-	2-Jul-56	NFRC 100137
Canada	-	Fairview	-	-	*Picea glauca*	M	-	16-Jul-60	NFRC 100117
Canada	-	63 McKenzie Highway	-	-	*Picea glauca*	F	-	22-Jul-54	NFRC 100122
Canada	-	Lac La Bonnet	-	-	*Picea glauca*	M	-	20-Jun-52	NFRC 100180
Germany	Saxony	Rachlan	51.133	14.517	*Picea excelsa*	M	Schrifze	-	UASM 3321
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	M	C. MacQuarrie	20-Jun-05	D56
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	M	C. MacQuarrie	20-Jun-05	D57
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	M	C. MacQuarrie	20-Jun-05	D58
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	M	C. MacQuarrie	20-Jun-05	D59
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	M	C. MacQuarrie	20-Jun-05	D60
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	F	C. MacQuarrie	20-Jun-05	D61
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	F	C. MacQuarrie	20-Jun-05	D62
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	F	C. MacQuarrie	20-Jun-05	D63
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	F	C. MacQuarrie	20-Jun-05	D64
USA	Alaska	Anchorage	61.131	-149.540	*Picea pungens*	F	C. MacQuarrie	20-Jun-05	D65
USA	California	Tehachapi Mtn. Prk. 8 km S Tehachapi	35.069	-118.484	*Abies concolour*	M	J.A. DeBenedictis & S. Meredith	16-Jun-80	EMEC 70 298
USA	California	Tehachapi Mtn. Prk. 8 km S Tehachapi	35.069	-118.484	*Abies concolour*	F	J.A. DeBenedictis & S. Meredith	16-Jun-80	EMEC 70 299
USA	California	Tehachapi Mtn. Prk. 8 km S Tehachapi	35.069	-118.484	*Abies concolour*	M	J.A. DeBenedictis & S. Meredith	16-Jun-80	EMEC 70 307
USA	California	Tehachapi Mtn. Prk.	35.069	-118.484	*Abies concolour*	M	J.A. Debenedictis & J.A. Powell	16-Jun-81	EMEC 70 308
USA	California	Tehachapi Mtn. Prk. 8 air km SW Tehachapi	35.069	-118.484	unknown	F	J. Powell, D. Wagner & J. DeBenedictis	18-Jul-83	EMEC 70 312
USA	California	Tehachapi Mtn. Prk. 8 air km SW Tehachapi	35.069	-118.484	unknown	F	J. Powell, D. Wagner & J. DeBenedictis	18-Jul-83	EMEC 70 316
USA	California	Felton	37.051	-122.073	unknown	F	J. Powell	21-Jul-95	EMEC 70 370
USA	California	Madeline, 14 mi NE (USFS block IX)	41.051	-120.476	*Abies concolour*	F	J. Powell	1-Jul-74	EMEC 70 341
USA	California	Madeline, 14 mi NE (USFS block IX)	41.051	-120.476	*Abies concolour*	F	J. Powell	1-Jul-74	EMEC 70 372
USA	California	Adin, 9 mi NE (USFS block XI)	41.194	-120.945	*Abies concolour*	F	J. Powell	24-Jun-74	EMEC 70 337
USA	California	Cedar Pass, 12 mi S (USFS block VI)	41.562	-120.269	*Abies concolour*	F	J. Powell	Jun-74	EMEC 70 328
USA	California	Cedar Pass, 10 mi N (USFS block II)	41.562	-120.269	*Abies concolour*	F	J. Powell	30-Jun-74	EMEC 70 330
USA	California	Cedar Pass, 10 mi N (USFS block II)	41.562	-120.269	*Abies concolour*	F	J. Powell	30-Jun-74	EMEC 70 332
USA	California	Cedar Pass, 12 mi S (USFS block VI)	41.562	-120.269	*Abies concolour*	F	J. Powell	Jun-74	EMEC 70 333
USA	California	Cedar Pass, 12 mi S (USFS block VI)	41.562	-120.269	*Abies concolour*	F	J. Powell	26-Jun-74	EMEC 70 334

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #
USA	California	Cedar Pass, 12 mi S (USFS block VI)	41.562	-120.269	*Abies concolour*	F	J. Powell	Jun-74	EMEC 70 335
USA	California	Cedar Pass, 3 mi S (USFS block V)	41.562	-120.269	*Abies concolour*	F	J. Powell	6-Jul-74	EMEC 70 336
USA	California	Cedar Pass 6 mi NW (USFS block III)	41.562	-120.269	*Abies concolour*	F	J. Powell	7-Jul-74	EMEC 70 338
USA	California	Cedar Pass 6 mi NW (USFS block III)	41.562	-120.269	*Abies concolour*	F	J. Powell	7-Jul-74	EMEC 70 389
USA	California	Willow Cr. 10 km SW Cazule	41.661	-122.508	*Pseudotsuga menziesii*	F	J. Powell	24-Jun-80	EMEC 70 339
USA	California	Prather Cr. 7 km SW McDoel	41.834	-122.066	*Pseudotsuga menziesii*	F	J.A. Debenedictis & J.A. Powell	25-Jun-80	EMEC 70 340
USA	California	Fandango Prk. 2 mi. NW (USFS block 1)	41.850	-120.272	*Abies concolour*	F	J. Powell	26-Jun-74	EMEC 70 324
USA	California	Fandango Prk. 2 mi. NW (USFS block 1)	41.850	-120.272	*Abies concolour*	F	J. Powell	26-Jun-74	EMEC 70 325
USA	California	Fandango Prk. 2 mi. NW (USFS block 1)	41.850	-120.272	*Abies concolour*	F	J. Powell	26-Jun-74	EMEC 70 326
USA	California	Fandango Prk. 2 mi. NW (USFS block 1)	41.850	-120.272	*Abies concolour*	F	J. Powell	26-Jun-74	EMEC 70 327
USA	California	Fandango Prk. 2 mi. NW (USFS block 1)	41.850	-120.272	*Abies concolour*	M	J. Powell	26-Jun-74	EMEC 70 329
USA	California	Fandango Prk. 2 mi. NW (USFS block 1)	41.850	-120.272	*Abies concolour*	F	J. Powell	24-Jun-74	EMEC 70 385
USA	California	Fandango Prk. 2 mi. NW (USFS block 1)	41.850	-120.272	*Abies concolour*	M	J. Powell	26-Jun-74	EMEC 70 386
USA	California	Masonite Rd.	-	-	*Pseudotsuga menziesii*	M	W.J.A. Volney	21-May-81	EMEC 70 350
USA	California	Masonite Rd.	-	-	*Pseudotsuga menziesii*	F	W.J.A. Volney	21-May-81	EMEC 70 351
USA	California	Masonite Rd.	-	-	*Pseudotsuga menziesii*	F	W.J.A. Volney	20-May-81	EMEC 70 356
USA	Idaho	Targhee	44.500	-111.334	*Pseudotsuga menziesii*	M	J. Weatherby	11-Jul-86	EMEC 70 352
USA	Idaho	Targhee	44.500	-111.334	*Pseudotsuga menziesii*	M	J. Weatherby	11-Jul-86	EMEC 70 353
USA	Idaho	Targhee	44.500	-111.334	*Pseudotsuga menziesii*	F	J. Weatherby	11-Jul-86	EMEC 70 354
USA	Idaho	4th of July Creek, N. of Salmon	45.176	-113.896	unknown	F	Margot May	4-Aug-65	UASM 3296
USA	Montana	Birch Creek 15 mi NW Dillon	44.639	-112.612	*Pseudotsuga menziesii*	F	D.C. Fellin	-	UASM 3457
USA	Montana	Bear Creek G.S. Cameron, Budworm Plot 3	45.203	-111.679	*Picea engelmanni*	F	H.R. Dodge	-	UASM 3474
USA	Montana	Bear Creek G.S. Cameron, Budworm Plot 3	45.203	-111.679	unknown	F	H.R. Dodge	-	UASM 3476
USA	Montana	Bear Creek G.S. Cameron, Budworm Plot 3	45.203	-111.679	*Picea engelmanni*	F	H.R. Dodge	-	UASM 3495
USA	Montana	Bear Creek G.S. Cameron, Budworm Plot 3	45.203	-111.679	*Picea engelmanni*	F	H.R. Dodge	-	UASM 3496
USA	Montana	Gallatin Valley	45.675	-111.012	unknown	M	H.R. Dodge	-	UASM 3450
USA	Montana	Gallatin Valley	45.675	-111.012	unknown	M	H.R. Dodge	-	UASM 3453
USA	Montana	3 mi SE Pipestone Pass	45.857	-112.43°	*Pseudotsuga menziesii*	F	D.C. Fellin	-	UASM 3463
USA	Montana	Flathead G.S. Wilsall	45.994	-110.660	*Pseudotsuga menziesii*	-	H.R. Dodge	-	UASM 3513
USA	Montana	Dry Range	45.800	-111.284	*Pinus ponderosa*	F	H.R. Dodge	-	UASM 3512
USA	Montana	western	-	-	unknown	M	-	-	UASM 3508
USA	Montana	western	-	-	unknown	F	-	-	UASM 3509
USA	Montana	western	-	-	unknown	M	-	-	UASM 3511
USA	Oregon	Burns	43.586	-119.054	*Pseudotsuga menziesii*	M	-	-	UASM 3767

Country	Prov./State	Location	Dec. Lat	Dec. Long	Host	Sex	Collector	Collection Date	Specimen #*
USA	Oregon	Burns	43.586	-119.054	*Pseudotsuga menziesii*	M	-	-	UASM 3768
USA	Oregon	Burns	43.586	-119.054	*Abies grandis*	M	-	-	UASM 3783
USA	Oregon	Burns	43.586	-119.054	*Abies grandis*	M	-	-	UASM 3786
USA	Oregon	Corvallis	44.335	-123.154	unknown	F	J. Adams	25-Jul-01	AR104
USA	Oregon	Corvallis	44.335	-123.154	unknown	M	J. Adams	25-Jul-01	AR105
USA	Oregon	Corvallis	44.335	-123.154	unknown	F	J. Adams	25-Jul-01	AR106
USA	Oregon	Corvallis	44.335	-123.154	unknown	F	J. Adams	25-Jul-01	AR107
USA	Oregon	Philomath, 12 km SW	44.540	-123.368	unknown	M	M.D. Shorb	-	UASM 3435
USA	Oregon	Philomath, 12 km SW	44.540	-123.368	unknown	M	M.D. Shorb	-	UASM 3438
USA	Oregon	Philomath, 12 km SW	44.540	-123.368	unknown	M	J.K. McPike	-	UASM 3446
USA	Oregon	Philomath, 12 km SW	44.540	-123.368	unknown	M	M.D. Shorb	-	UASM 3728
USA	Oregon	Philomath, 12 km SW	44.540	-123.368	unknown	M	M.D. Shorb	-	UASM 3730
USA	Oregon	Philomath, 12 km SW	44.540	-123.368	unknown	M	J. K. McPike	-	UASM 3754
USA	Oregon	Philomath, Blakesley Creek	44.601	-123.383	unknown	F	AVZ Brower & DD Judd	19-Aug-04	UASM 3429
USA	Oregon	Philomath, Blakesley Creek	44.601	-123.383	unknown	F	AVZ Brower & DD Judd	11-Jul-04	UASM 3430
USA	Oregon	Vic. Gronnel Rd. Ca. 2 mi, Elsie	45.866	-123.595	unknown	M	S. G. Jewett Jr.	7-Aug-71	UASM 3431
USA	Oregon	Vic. Gronnel Rd. Ca. 2 mi E Elsie	45.866	-123.595	unknown	F	S.G. Jewett Jr.	7-Aug-71	UASM 3644
USA	Washington	7 mi NW Conconully	48.55.737 10	-119.750	*Pseudotsuga menziesii*	M	J. Powell	3-Jul-88	EMEC 70 183

* Description of specimen identifiers: AR#= A. Roe DNA number set 1, D# = A. Roe DNA number set 2; NFRC = Northern Forestry Research Center; UASM = University of Alberta Strickland Museum; EMEC= Essig Museum Entomology Collection.

Biography

On November 30th, 1978 I was born in a small rural hospital in Vulcan, Alberta, and I spent the next eighteen glorious years getting dirty on a family farm homesteaded by my great grandfather. Growing up on a farm gave me the luxury of exploring to my heart's content. I had a voracious appetite for natural history, and asked questions about everything. My parents always encouraged and supported my love of nature despite a plethora of pets and the occasional escaped snake (sorry Mom!).

Throughout school my fascination with natural history continued and I developed an aptitude for science, which was encouraged by several memorable science teachers. By sheer luck and a farsighted biology teacher, the Vulcan high school was able to offer an after-school marine biology program. This was exceptionally odd for a land-locked rural school, but it allowed budding scientists like myself to delve into the incredible world of marine life, navigation, and scientific inquiry. I spent a week every year for four years sailing around the San Juan Islands off the coast of Victoria with my eyes glued to binoculars and fingers deep in tide pools. It was an incredible experience and has certainly helped make me what I am today.

Immediately after high school, I enrolled in the Environmental Biology program at the University of Alberta and soon fell into a crowd of like-mind peers, something that I did not have in my rural upbringing. My undergraduate program consisted of an eclectic mix of courses ranging from animal community ecology to fungal biology, however insect taxonomy truly piqued my interest. Actually, it was while building my insect collection for the course that I had an epiphany. I realized that I wanted to devote the rest of my life to the study of insects. I got my foot in the door doing an honors thesis with Dr. Felix Sperling on the *Dioryctria* of Alberta. I turned this project into a Masters program, which was expanded into my Doctoral thesis, after a great deal of thought and consideration.

I have had many incredible experiences over the course of my graduate program. I have had the opportunity to travel and collect all over the world, from the high Arctic to the Great Barrier Reef off the coast of Australia. Course work and discussions have helped me to expand my intellectual horizons and participate in a number of fascinating

debates. Of course, graduate school would have been nothing without the friends and peers that have helped me along the way and challenged the way I think. These past five years have been among the best of my life, and the experiences and friends I have had will continue to influence me for the rest of my life.